

! VH1

```

! 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80
agg gtc acc atg acc agg gac acg tcc atc agc aca gcc tac atg
5 ! 81 82 82a 82b 82c 83 84 85 86 87 88 89 90 91 92
gag ctg agc agg ctg aga tct gac gac acg gcc gtg tat tac tgt
! 93 94 95
gcg aga ga ! 1-02# 1
aga gtc acc att acc agg gac aca tcc gcg agc aca gcc tac atg
10 gag ctg agc agc ctg aga tct gaa gac acg gct gtg tat tac tgt
gcg aga ga ! 1-03# 2
aga gtc acc atg acc agg aac acc tcc ata agc aca gcc tac atg
gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt
gcg aga gg ! 1-08# 3
15 aga gtc acc atg acc aca gac aca tcc acg agc aca gcc tac atg
gag ctg agg agc ctg aga tct gac gac acg gcc gtg tat tac tgt
gcg aga ga ! 1-18# 4
aga gtc acc atg acc gag gac aca tct aca gac aca gcc tac atg
gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt
20 gca aca ga ! 1-24# 5
aga gtc acc att acc agg gac agg tct atg agc aca gcc tac atg
gag ctg agc agc ctg aga tct gag gac aca gcc atg tat tac tgt
gca aga ta ! 1-45# 6
aga gtc acc atg acc agg gac acg tcc acg agc aca gtc tac atg
25 gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt
gcg aga ga ! 1-46# 7
aga gtc acc att acc agg gac atg tcc aca agc aca gcc tac atg
gag ctg agc agc ctg aga tcc gag gac acg gcc gtg tat tac tgt
gcg gca ga ! 1-58# 8
30 aga gtc acg att acc gcg gac gaa tcc acg agc aca gcc tac atg
gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt
gcg aga ga ! 1-69# 9
aga gtc acg att acc gcg gac aaa tcc acg agc aca gcc tac atg
gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt
35 gcg aga ga ! 1-e# 10.
aga gtc acc ata acc gcg gac acg tct aca gac aca gcc tac atg
gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt
gca aca ga ! 1-f# 11

```

agg ctc acc atc acc aag gac acc tcc aaa aac cag gtg gtc ctt  
aca atg acc aac atg gac cct gtg gac aca gcc aca tat tac tgt  
gca cac aga c! 2-05# 12

5 agg ctc acc atc tcc aag gac acc tcc aaa agc cag gtg gtc ctt  
acc atg acc aac atg gac cct gtg gac aca gcc aca tat tac tgt  
gca cgg ata c! 2-26# 13

agg ctc acc atc tcc aag gac acc tcc aaa aac cag gtg gtc ctt  
aca atg acc aac atg gac cct gtg gac aca gcc acg tat tac tgt

10 gca cgg ata c! 2-70# 14

! VH3

cga ttc acc atc tcc aga gac aac gcc aag aac tca ctg tat ctg  
caa atg aac agc ctg aga gcc gag gac acg gct gtg tat tac tgt  
gcg aga ga ! 3-07# 15

15 cga ttc acc atc tcc aga gac aac gcc aag aac tcc ctg tat ctg  
caa atg aac agt ctg aga gct gag gac acg gcc ttg tat tac tgt  
gca aaa gat a! 3-09#16

cga ttc acc atc tcc agg gac aac gcc aag aac tca ctg tat ctg  
caa atg aac agc ctg aga gcc gag gac acg gcc gtg tat tac tgt

20 gcg aga ga ! 3-11# 17

cga ttc acc atc tcc aga gaa aat gcc aag aac tcc ttg tat ctt  
caa atg aac agc ctg aga gcc ggg gac acg gct gtg tat tac tgt  
gca aga ga ! 3-13# 18

aga ttc acc atc tca aga gat gat tca aaa aac acg ctg tat ctg

25 caa atg aac agc ctg aaa acc gag gac aca gcc gtg tat tac tgt  
acc aca ga ! 3-15# 19

cga ttc acc atc tcc aga gac aac gcc aag aac tcc ctg tat ctg  
caa atg aac agt ctg aga gcc gag gac acg gcc ttg tat cac tgt  
gcg aga ga ! 3-20# 20

30 cga ttc acc atc tcc aga gac aac gcc aag aac tca ctg tat ctg  
caa atg aac agc ctg aga gcc gag gac acg gct gtg tat tac tgt  
gcg aga ga ! 3-21# 21

cgg ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg  
caa atg aac agc ctg aga gcc gag gac acg gcc gta tat tac tgt

35 gcg aaa ga ! 3-23# 22

cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg  
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt  
gcg aaa ga ! 3-30# 23

cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg

40 caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt

gcg aga ga ! 3303# 24  
cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg  
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt  
gcg aaa ga ! 3305# 25  
5 cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg  
caa atg aac agc ctg aga gcc gag gac acg gct gtg tat tac tgt  
gcg aga ga ! 3-33# 26  
cga ttc acc atc tcc aga gac aac agc aaa aac tcc ctg tat ctg  
caa atg aac agt ctg aga act gag gac acc gcc ttg tat tac tgt  
10 gca aaa gat a! 3-43#27  
cga ttc acc atc tcc aga gac aat gcc aag aac tca ctg tat ctg  
caa atg aac agc ctg aga gac gag gac acg gct gtg tat tac tgt  
gcg aga ga ! 3-48# 28  
aga ttc acc atc tca aga gat ggt tcc aaa agc atc gcc tat ctg  
15 caa atg aac agc ctg aaa acc gag gac aca gcc gtg tat tac tgt  
act aga ga ! 3-49# 29  
cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctt  
caa atg aac agc ctg aga gcc gag gac acg gcc gtg tat tac tgt  
gcg aga ga ! 3-53# 30  
20 aga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctt  
caa atg gcc agc ctg aga gct gag gac atg gct gtg tat tac tgt  
gcg aga ga ! 3-64# 31  
aga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctt  
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt  
25 gcg aga ga ! 3-66# 32  
aga ttc acc atc tca aga gat gat tca aag aac tca ctg tat ctg  
caa atg aac agc ctg aaa acc gag gac acg gcc gtg tat tac tgt  
gct aga ga ! 3-72# 33  
agg ttc acc atc tcc aga gat gat tca aag aac acg gcg tat ctg  
30 caa atg aac agc ctg aaa acc gag gac acg gcc gtg tat tac tgt  
act aga ca ! 3-73# 34  
cga ttc acc atc tcc aga gac aac gcc aag aac acg ctg tat ctg  
caa atg aac agt ctg aga gcc gag gac acg gct gtg tat tac tgt  
gca aga ga ! 3-74# 35  
35 aga ttc acc atc tcc aga gac aat tcc aag aac acg ctg cat ctt  
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt  
aag aaa ga ! 3-d# 36  
! VH4  
40 cga gtc acc ata tca gta gac aag tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gcc gcg gac acg gcc gtg tat tac tgt

gcg aga ga ! 4-04# 37  
cga gtc acc atg tca gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gcc gtg gac acg gcc gtg tat tac tgt  
gcg aga aa ! 4-28# 38  
5 cga gtt acc ata tca gta gac acg tct aag aac cag ttc tcc ctg  
aag ctg agc tct gtg act gcc gcg gac acg gcc gtg tat tac tgt  
gcg aga ga ! 4301# 39  
cga gtc acc ata tca gta gac agg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gcc gcg gac acg gcc gtg tat tac tgt  
10 gcc aga ga ! 4302# 40  
cga gtt acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg act gcc gca gac acg gcc gtg tat tac tgt  
gcc aga ga ! 4304# 41  
cga gtt acc ata tca gta gac acg tct aag aac cag ttc tcc ctg  
15 aag ctg agc tct gtg act gcc gcg gac acg gcc gtg tat tac tgt  
gcg aga ga ! 4-31# 42  
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gcc gcg gac acg gct gtg tat tac tgt  
gcg aga ga ! 4-34# 43  
20 cga gtc acc ata tcc gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gcc gca gac acg gct gtg tat tac tgt  
gcg aga ca ! 4-39# 44  
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gct gcg gac acg gcc gtg tat tac tgt  
25 gcg aga ga ! 4-59# 45  
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg  
aag ctg agc tct gtg acc gct gcg gac acg gcc gtg tat tac tgt  
gcg aga ga ! 4-61# 46  
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg  
30 aag ctg agc tct gtg acc gcc gca gac acg gcc gtg tat tac tgt  
gcg aga ga ! 4-b# 47  
! VH5  
cag gtc acc atc tca gcc gac aag tcc atc agc acc gcc tac ctg  
cag tgg agc agc ctg aag gcc tcg gac acc gcc atg tat tac tgt  
35 gcg aga ca ! 5-51# 48  
cac gtc acc atc tca gct gac aag tcc atc agc act gcc tac ctg  
cag tgg agc agc ctg aag gcc tcg gac acc gcc atg tat tac tgt  
gcg aga ! 5-a# 49  
! VH6  
40 cga ata acc atc aac cca gac aca tcc aag aac cag ttc tcc ctg





Table 2: Enzymes that either cut 15 or more human GLGs or have 5+-base recognition in FR3

Typical entry:

REname Recognition #sites  
GLGid#:base# GLGid#:base# GLGid#:base#.....

5

BstEII Ggtacc 2

1: 3 48: 3

There are 2 hits at base# 3

10

MaeIII gtnac 36

1: 4 2: 4 3: 4 4: 4 5: 4 6: 4

7: 4 8: 4 9: 4 10: 4 11: 4 37: 4

37: 58 38: 4 38: 58 39: 4 39: 58 40: 4

40: 58 41: 4 41: 58 42: 4 42: 58 43: 4

15

43: 58 44: 4 44: 58 45: 4 45: 58 46: 4

46: 58 47: 4 47: 58 48: 4 49: 4 50: 58

There are 24 hits at base# 4

Tsp45I gtsac 33

20

1: 4 2: 4 3: 4 4: 4 5: 4 6: 4

7: 4 8: 4 9: 4 10: 4 11: 4 37: 4

37: 58 38: 4 38: 58 39: 58 40: 4 40: 58

41: 58 42: 58 43: 4 43: 58 44: 4 44: 58

45: 4 45: 58 46: 4 46: 58 47: 4 47: 58

25

48: 4 49: 4 50: 58

There are 21 hits at base# 4

HphI tcacc 45

30

1: 5 2: 5 3: 5 4: 5 5: 5 6: 5

7: 5 8: 5 11: 5 12: 5 12: 11 13: 5

14: 5 15: 5 16: 5 17: 5 18: 5 19: 5

20: 5 21: 5 22: 5 23: 5 24: 5 25: 5

26: 5 27: 5 28: 5 29: 5 30: 5 31: 5

32: 5 33: 5 34: 5 35: 5 36: 5 37: 5

35

38: 5 40: 5 43: 5 44: 5 45: 5 46: 5

47: 5 48: 5 49: 5

There are 44 hits at base# 5

NlaIII CATG

26

1: 9 1: 42 2: 42 3: 9 3: 42 4: 9  
4: 42 5: 9 5: 42 6: 42 6: 78 7: 9  
7: 42 8: 21 8: 42 9: 42 10: 42 11: 42  
5 12: 57 13: 48 13: 57 14: 57 31: 72 38: 9  
48: 78 49: 78

There are 11 hits at base# 42

There are 1 hits at base# 48 Could cause raggedness.

10 BsaJI Ccnngg

37

1: 14 2: 14 5: 14 6: 14 7: 14 8: 14  
8: 65 9: 14 10: 14 11: 14 12: 14 13: 14  
14: 14 15: 65 17: 14 17: 65 18: 65 19: 65  
20: 65 21: 65 22: 65 26: 65 29: 65 30: 65  
15 33: 65 34: 65 35: 65 37: 65 38: 65 39: 65  
40: 65 42: 65 43: 65 48: 65 49: 65 50: 65  
51: 14

There are 23 hits at base# 65

There are 14 hits at base# 14

20

AluI AGct

42

1: 47 2: 47 3: 47 4: 47 5: 47 6: 47  
7: 47 8: 47 9: 47 10: 47 11: 47 16: 63  
23: 63 24: 63 25: 63 31: 63 32: 63 36: 63  
25 37: 47 37: 52 38: 47 38: 52 39: 47 39: 52  
40: 47 40: 52 41: 47 41: 52 42: 47 42: 52  
43: 47 43: 52 44: 47 44: 52 45: 47 45: 52  
46: 47 46: 52 47: 47 47: 52 49: 15 50: 47

There are 23 hits at base# 47

30 There are 11 hits at base# 52 Only 5 bases from 47

BlpI GCtnagc

21

1: 48 2: 48 3: 48 5: 48 6: 48 7: 48  
8: 48 9: 48 10: 48 11: 48 37: 48 38: 48  
35 39: 48 40: 48 41: 48 42: 48 43: 48 44: 48  
45: 48 46: 48 47: 48

There are 21 hits at base# 48



Hpy188I TCNga 17  
 1: 64 2: 64 3: 64 4: 64 5: 64 6: 64  
 7: 64 8: 64 9: 64 10: 64 11: 64 16: 57  
 20: 57 27: 57 35: 57 48: 67 49: 67

5 There are 11 hits at base# 64  
 There are 4 hits at base# 57  
 There are 2 hits at base# 67 Could be ragged.

MslI CAYNNnnRTG 44  
 10 1: 72 2: 72 3: 72 4: 72 5: 72 6: 72  
 7: 72 8: 72 9: 72 10: 72 11: 72 15: 72  
 17: 72 18: 72 19: 72 21: 72 23: 72 24: 72  
 25: 72 26: 72 28: 72 29: 72 30: 72 31: 72  
 32: 72 33: 72 34: 72 35: 72 36: 72 37: 72  
 15 38: 72 39: 72 40: 72 41: 72 42: 72 43: 72  
 44: 72 45: 72 46: 72 47: 72 48: 72 49: 72  
 50: 72 51: 72

There are 44 hits at base# 72

20 BsiEI CGRYcg 23  
 1: 74 3: 74 4: 74 5: 74 7: 74 8: 74  
 9: 74 10: 74 11: 74 17: 74 22: 74 30: 74  
 33: 74 34: 74 37: 74 38: 74 39: 74 40: 74  
 41: 74 42: 74 45: 74 46: 74 47: 74

25 There are 23 hits at base# 74

EaeI Yggccr 23  
 1: 74 3: 74 4: 74 5: 74 7: 74 8: 74  
 9: 74 10: 74 11: 74 17: 74 22: 74 30: 74  
 30 33: 74 34: 74 37: 74 38: 74 39: 74 40: 74  
 41: 74 42: 74 45: 74 46: 74 47: 74

There are 23 hits at base# 74

EagI Cggccg 23  
 35 1: 74 3: 74 4: 74 5: 74 7: 74 8: 74  
 9: 74 10: 74 11: 74 17: 74 22: 74 30: 74  
 33: 74 34: 74 37: 74 38: 74 39: 74 40: 74

41: 74 42: 74 45: 74 46: 74 47: 74

There are 23 hits at base# 74

HaeIII GGcc

27

5 1: 75 3: 75 4: 75 5: 75 7: 75 8: 75  
9: 75 10: 75 11: 75 16: 75 17: 75 20: 75  
22: 75 30: 75 33: 75 34: 75 37: 75 38: 75  
39: 75 40: 75 41: 75 42: 75 45: 75 46: 75  
47: 75 48: 63 49: 63

10 There are 25 hits at base# 75

Bst4CI ACNgt 65°C

63 Sites There is a third isoschimer

1 86 2: 86 3: 86 4: 86 5: 86 6: 86  
7: 34 7: 86 8: 86 9: 86 10: 86 11: 86  
15 12: 86 13: 86 14: 86 15: 36 15: 86 16: 53  
16: 86 17: 36 17: 86 18: 86 19: 86 20: 53  
20: 86 21: 36 21: 86 22: 0 22: 86 23: 86  
24: 86 25: 86 26: 86 27: 53 27: 86 28: 36  
28: 86 29: 86 30: 86 31: 86 32: 86 33: 36  
20 33: 86 34: 86 35: 53 35: 86 36: 86 37: 86  
38: 86 39: 86 40: 86 41: 86 42: 86 43: 86  
44: 86 45: 86 46: 86 47: 86 48: 86 49: 86  
50: 86 51: 0 51: 86

There are 51 hits at base# 86 All the other sites are well away

25

HpyCH4III ACNgt

63

1: 86 2: 86 3: 86 4: 86 5: 86 6: 86  
7: 34 7: 86 8: 86 9: 86 10: 86 11: 86  
12: 86 13: 86 14: 86 15: 36 15: 86 16: 53  
30 16: 86 17: 36 17: 86 18: 86 19: 86 20: 53  
20: 86 21: 36 21: 86 22: 0 22: 86 23: 86  
24: 86 25: 86 26: 86 27: 53 27: 86 28: 36  
28: 86 29: 86 30: 86 31: 86 32: 86 33: 36  
33: 86 34: 86 35: 53 35: 86 36: 86 37: 86  
35 38: 86 39: 86 40: 86 41: 86 42: 86 43: 86  
44: 86 45: 86 46: 86 47: 86 48: 86 49: 86  
50: 86 51: 0 51: 86

There are 51 hits at base# 86

HinFI Gantc 43

	2: 2	3: 2	4: 2	5: 2	6: 2	7: 2
5	8: 2	9: 2	9: 22	10: 2	11: 2	15: 2
	16: 2	17: 2	18: 2	19: 2	19: 22	20: 2
	21: 2	23: 2	24: 2	25: 2	26: 2	27: 2
	28: 2	29: 2	30: 2	31: 2	32: 2	33: 2
	33: 22	34: 22	35: 2	36: 2	37: 2	38: 2
10	40: 2	43: 2	44: 2	45: 2	46: 2	47: 2
	50: 60					

There are 38 hits at base# 2

MlyI GAGTCNNNNn 18

15	2: 2	3: 2	4: 2	5: 2	6: 2	7: 2
	8: 2	9: 2	10: 2	11: 2	37: 2	38: 2
	40: 2	43: 2	44: 2	45: 2	46: 2	47: 2

There are 18 hits at base# 2

PleI gagtc 18

	2: 2	3: 2	4: 2	5: 2	6: 2	7: 2
	8: 2	9: 2	10: 2	11: 2	37: 2	38: 2
	40: 2	43: 2	44: 2	45: 2	46: 2	47: 2

There are 18 hits at base# 2

AciI Ccgc 24

25	2: 26	9: 14	10: 14	11: 14	27: 74	37: 62
	37: 65	38: 62	39: 65	40: 62	40: 65	41: 65
	42: 65	43: 62	43: 65	44: 62	44: 65	45: 62
	46: 62	47: 62	47: 65	48: 35	48: 74	49: 74

30 There are 8 hits at base# 62

There are 8 hits at base# 65

There are 3 hits at base# 14

There are 3 hits at base# 74

There are 1 hits at base# 26

35 There are 1 hits at base# 35

--- Gcgg 11

	8: 91	9: 16	10: 16	11: 16	37: 67	39: 67
--	-------	-------	--------	--------	--------	--------





46: 64 48: 53 49: 53 50: 45 51: 53

There are 13 hits at base# 53

MnlI gagg

34

5 3: 67 3: 95 4: 51 5: 16 5: 67 6: 67  
 7: 67 8: 67 9: 67 10: 67 11: 67 15: 67  
 16: 67 17: 67 19: 67 20: 67 21: 67 22: 67  
 23: 67 24: 67 25: 67 26: 67 27: 67 28: 67  
 29: 67 30: 67 31: 67 32: 67 33: 67 34: 67

10 35: 67 36: 67 50: 67 51: 67

There are 31 hits at base# 67

HpyCH4V TGca

34

5: 90 6: 90 11: 90 12: 90 13: 90 14: 90  
 15 15: 44 16: 44 16: 90 17: 44 18: 90 19: 44  
 20: 44 21: 44 22: 44 23: 44 24: 44 25: 44  
 26: 44 27: 44 27: 90 28: 44 29: 44 33: 44  
 34: 44 35: 44 35: 90 36: 38 48: 44 49: 44  
 50: 44 50: 90 51: 44 51: 52

20 There are 21 hits at base# 44

There are 1 hits at base# 52

AccI GTmkac

13

5-base recognition

7: 37 11: 24 37: 16 38: 16 39: 16 40: 16  
 25 41: 16 42: 16 43: 16 44: 16 45: 16 46: 16  
 47: 16

There are 11 hits at base# 16

SacII CCGCgg

8

6-base recognition

30 9: 14 10: 14 11: 14 37: 65 39: 65 40: 65  
 42: 65 43: 65

There are 5 hits at base# 65

There are 3 hits at base# 14

35 TfiI Gawtc

24

9: 22 15: 2 16: 2 17: 2 18: 2 19: 2  
 19: 22 20: 2 21: 2 23: 2 24: 2 25: 2

26: 2 27: 2 28: 2 29: 2 30: 2 31: 2  
32: 2 33: 2 33: 22 34: 22 35: 2 36: 2

There are 20 hits at base# 2

5 BsmAI Nnnnnngagac 19  
15: 11 16: 11 20: 11 21: 11 22: 11 23: 11  
24: 11 25: 11 26: 11 27: 11 28: 11 28: 56  
30: 11 31: 11 32: 11 35: 11 36: 11 44: 87  
48: 87

10 There are 16 hits at base# 11

BpmI ctccag 19  
15: 12 16: 12 17: 12 18: 12 20: 12 21: 12  
22: 12 23: 12 24: 12 25: 12 26: 12 27: 12  
15 28: 12 30: 12 31: 12 32: 12 34: 12 35: 12  
36: 12

There are 19 hits at base# 12

XmnI GAANNnnttc 12  
20 37: 30 38: 30 39: 30 40: 30 41: 30 42: 30  
43: 30 44: 30 45: 30 46: 30 47: 30 50: 30

There are 12 hits at base# 30

BsrI NCcagt 12  
25 37: 32 38: 32 39: 32 40: 32 41: 32 42: 32  
43: 32 44: 32 45: 32 46: 32 47: 32 50: 32

There are 12 hits at base# 32

BanII GRGCYc 11  
30 37: 51 38: 51 39: 51 40: 51 41: 51 42: 51  
43: 51 44: 51 45: 51 46: 51 47: 51

There are 11 hits at base# 51

Ecl136I GAGctc 11  
35 37: 51 38: 51 39: 51 40: 51 41: 51 42: 51  
43: 51 44: 51 45: 51 46: 51 47: 51

There are 11 hits at base# 51

SacI GAGCTc

11

37: 51    38: 51    39: 51    40: 51    41: 51    42: 51

43: 51      44: 51      45: 51      46: 51      47: 51

There are 11 hits at base# 51

Table 3: Synthetic 3-23 FR3 of human heavy chains showing positions of possible cleavage sites

```

! Sites engineered into the synthetic gene are shown in upper case
DNA
! with the RE name between vertical bars (as in | XbaI |).
5 ! RERSs frequently found in GLGs are shown below the synthetic
sequence
! with the name to the right (as in gtn ac=MaeIII(24), indicating
that
! 24 of the 51 GLGs contain the site).
10 !
! |---FR3---
! 89 90 (codon
# in
! R F
15 synthetic 3-23)
|cgc|ttc| 6
! Allowed DNA |cgn|tty|
! |agr|
! ga ntc =
20 HinfI(38)
! ga gtc =
PleI(18)
! ga wtc =
TfiI(20)
25 ! gtn ac =
MaeIII(24)
! gts ac =
Tsp45I(21)
! tc acc =
30 HphI(44)
!
! -----FR3-----
! 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
! T I S R D N S K N T L Y L Q M
35 |act|atc|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|
51
!allowed|acn|ath|tcn|cgn|gay|aay|tcn|aar|aay|acn|ttr|tay|ttr|car|atg|
! |agy|agr| |agy| |ctn| |ctn|
! | ga|gac = BsmAI(16) ag ct =
40 AluI(23)

```

```

!           c|tcc ag = BpmI(19)           g ctn agc =
BlpI(21)
!           |           |           g aan nnn ttc = XmnI(12)
!           | XbaI |           tg ca =
5 HpyCH4V(21)
!
!           ---FR3----->|
!           106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
!           N   S   L   R   A   E   D   T   A   V   Y   Y   C   A   K
10 |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgc|gct|aaa| 96
!allowed|aay|tcn|ttr|cgn|gcn|gar|gay|acn|gcn|gtt|tay|tay|tgy|gcn|aar|
!           |agy|ctn|agr|           |           |
!           |           | cc nng g = BsaJI(23)           ac ngt = Bst4CI(51)
!           |           | aga tct = BglII(10)           |           ac ngt =
15 HpyCH4III(51)
!           |           | Rga tcY = BstYI(11)           |           ac ngt = TaaI(51)
!           |           |           c ayn nnn rtc = MslI(44)
!           |           |           cg ryc g = BsiEI(23)
!           |           |           yg gcc r = EaeI(23)
20 !           |           |           cg gcc g = EagI(23)
!           |           |           |g gcc = HaeIII(25)
!           |           |           gag g = MnlI(31)|
!           |AflII |           | PstI |

```

Table 4: REaptors, Extenders, and Bridges used for Cleavage and Capture of Human Heavy Chains in FR3.

**A: HpyCH4V Probes of actual human HC genes**

!HpyCH4V in FR3 of human HC, bases 35-56; only those with TGca site

5	TGca;10,		
	RE recognition:tgca		of length 4 is expected at
10			
	1	6-1	agttctccctgcagctgaactc
	2	3-11,3-07,3-21,3-72,3-48	cactgtatctgcaaatgaacag
10	3	3-09,3-43,3-20	ccctgtatctgcaaatgaacag
	4	5-51	ccgcctacctgcagtgaggagcag
	5	3-15,3-30,3-30.5,3-30.3,3-74,3-23,3-33	cgctgtatctgcaaatgaacag
	6	7-4.1	cggcatatctgcagatctgcag
	7	3-73	cggcgtatctgcaaatgaacag
15	8	5-a	ctgcctacctgcagtgaggagcag
	9	3-49	tcgcctatctgcaaatgaacag

**B: HpyCH4V REaptors, Extenders, and Bridges**

**B.1 REaptors**

! Cutting HC lower strand:

20 ! TmKeller for 100 mM NaCl, zero formamide

! Eaptors for cleavage

		$T_m^W$	$T_m^K$
	(ON_HCFR36-1) 5'-agttctcccTGCAgctgaactc-3'	68.0	64.5
	(ON_HCFR36-1A) 5'-ttctcccTGCAgctgaactc-3'	62.0	62.5
	(ON_HCFR36-1B) 5'-ttctcccTGCAgctgaac-3'	56.0	59.9
25	(ON_HCFR33-15) 5'-cgctgtatcTGCAaatgaacag-3'	64.0	60.8
	(ON_HCFR33-15A) 5'-ctgtatcTGCAaatgaacag-3'	56.0	56.3
	(ON_HCFR33-15B) 5'-ctgtatcTGCAaatgaac-3'	50.0	53.1
	(ON_HCFR33-11) 5'-cactgtatcTGCAaatgaacag-3'	62.0	58.9
	(ON_HCFR35-51) 5'-ccgcctaccTGCAgtggagcag-3'	74.0	70.1

30 !

**B.2 Segment of synthetic 3-23 gene into which captured CDR3 is to be cloned**

! XbaI...

!D323\* cgCttcacTaag tcT aga gac aaC tctT aag aaT acT ctC taC

35 ! scab..... designed gene 3-23 gene.....

!

! HpyCH4V

! .. .. AflIII...

! Ttg caG atg aac agc TtA agG . . .  
!  
!

### B.3 Extender and Bridges

5 ! Extender (bottom strand):  
!  
(ON\_HCHpyEx01) 5'-cAAgTAgAgAgTATTcTTAgAgTTgTcTcTAGAcTTAgTgAAgcg-3'  
! ON\_HCHpyEx01 is the reverse complement of  
! 5'-cgCttcacTaag tcT aga gac aaC tcT aag aaT acT ctC taC Ttg -3'  
10 !  
! Bridges (top strand, 9-base overlap):  
!  
(ON\_HCHpyBr016-1) 5'-cgCttcacTaag tcT aga gac aaC tcT aag-  
aaT acT ctC taC Ttg CAgctgaac-3' {3'-term C is  
15 blocked}  
!  
! 3-15 et al. + 3-11  
(ON\_HCHpyBr023-15) 5'-cgCttcacTaag tcT aga gac aaC tcT aag-  
aaT acT ctC taC Ttg CAAatgaac-3' {3'-term C is  
20 blocked}  
!  
!  
! 5-51  
(ON\_HCHpyBr045-51) 5'-cgCttcacTaag tcT aga gac aaC tcT aag-  
aaT acT ctC taC Ttg CAgtgaggac-3' {3'-term C is  
25 blocked}  
!  
! PCR primer (top strand)  
!  
(ON\_HCHpyPCR) 5'-cgCttcacTaag tcT aga gac-3'  
30 !

---

### C: B1pI Probes from human HC GLGs

1	1-58,1-03,1-08,1-69,1-24,1-45,1-46,1-f,1-e
	acatggaGCTGAGCagcctgag
2	1-02
35	acatggaGCTGAGCaggctgag
3	1-18
	acatggagctgaggagcctgag

4 5-51,5-a  
acctgcagtggagcagcctgaa  
5 3-15,3-73,3-49,3-72  
atctgcaaataaacagcctgaa  
5 6 3303,3-33,3-07,3-11,3-30,3-21,3-23,3305,3-48  
atctgcaaataaacagcctgag  
7 3-20,3-74,3-09,3-43  
atctgcaaataaacagtctgag  
8 74.1  
10 atctgcagatctgcagcctaaa  
9 3-66,3-13,3-53,3-d  
atcttcaaataaacagcctgag  
10 3-64  
atcttcaaataaaggcagcctgag  
15 11 4301,4-28,4302,4-04,4304,4-31,4-34,4-39,4-59,4-61,4-b  
ccctgaaGCTGAGCtctgtgac  
12 6-1  
ccctgcagctgaactctgtgac  
13 2-70,2-05  
20 tccttacaatgaccaacatgga  
14 2-26  
tccttaccatgaccaacatgga

---

**D: B1pI REaptors, Extenders, and Bridges**

**D.1 REaptors**

		$T_m^W$	$T_m^K$
25	(BlpF3HC1-58) 5'-ac atg gaG CTG AGC agc ctg ag-3'	70	66.
			4
	(BlpF3HC6-1) 5'-cc ctg aag ctg agc tct gtg ac-3'	70	66.
			4
30	! BlpF3HC6-1 matches 4-30.1, not 6-1.		

**D.2 Segment of synthetic 3-23 gene into which captured CDR3 is to be cloned**

!  
B1pI  
35 ! XbaI...  
... ..  
!D323\* cgCttcacTaag TCT AGA gac aaC tcT aag aaT acT ctC taC Ttg  
caG atg aac



!  
!  
! AflIII...  
! agC TTA AGG

### D.3 Extender and Bridges

5 ! Bridges  
(BlpF3Br1) 5'-cgCttcacTcag tcT aga gaT aaC AGT aaA aaT acT TtG-  
taC Ttg caG Ctg a|GC agc ctg-3'  
(BlpF3Br2) 5'-cgCttcacTcag tcT aga gaT aaC AGT aaA aaT acT TtG-  
taC Ttg caG Ctg a|gc tct gtg-3'  
10 ! | lower strand is cut here  
! Extender  
(BlpF3Ext) 5'-  
TcAgcTgcAAgTAcAAAgTATTTTAcTgTTATcTcTAgAcTgAgTgAAgcg-3'  
! BlpF3Ext is the reverse complement of:  
15 ! 5'-cgCttcacTcag tcT aga gaT aaC AGT aaA aaT acT TtG taC Ttg caG  
Ctg a-3'  
!  
(BlpF3PCR) 5'-cgCttcacTcag tcT aga gaT aaC-3'

---

### E: *Hpy*CH4III Distinct GLG sequences surrounding site, bases 77-98

20	1	102#1,118#4,146#7,169#9,1e#10,311#17,353#30,404#37,4301
		ccgtgtattactgtgcgagaga
	2	103#2,307#15,321#21,3303#24,333#26,348#28,364#31,366#32
		ctgtgtattactgtgcgagaga
	3	108#3
25		ccgtgtattactgtgcgagagg
	4	124#5,1f#11
		ccgtgtattactgtgcaacaga
	5	145#6
		ccatgtattactgtgcaagata
30	6	158#8
		ccgtgtattactgtgcggcaga
	7	205#12
		ccacatattactgtgcacacag
	8	226#13
35		ccacatattactgtgcacggat
	9	270#14
		ccacgtattactgtgcacggat

	10	309#16,343#27
	ccttgattactgtgcaaaaga	
	11	313#18,374#35,61#50
	ctgtgtattactgtgcaagaga	
5	12	315#19
	ccgtgtattactgtaccacaga	
	13	320#20
	ccttgatcactgtgagagaga	
	14	323#22
10	ccgtatattactgtgcaaaaga	
	15	330#23,3305#25
	ctgtgtattactgtgcaaaaga	
	16	349#29
	ccgtgtattactgtactagaga	
15	17	372#33
	ccgtgtattactgtgctagaga	
	18	373#34
	ccgtgtattactgtactagaca	
	19	3d#36
20	ctgtgtattactgtgcaaaaga	
	20	428#38
	ccgtgtattactgtgagagaaa	
	21	4302#40,4304#41
	ccgtgtattactgtgccagaga	
25	22	439#44
	ctgtgtattactgtgagagaca	
	23	551#48
	ccatgtattactgtgagagaca	
	24	5a#49
30	ccatgtattactgtgagaga	

# **F: HpyCH4III REaptors, Extenders, and Bridges**

## **F.1 REaptors**

! ONs for cleavage of HC(lower) in FR3(bases 77-97)

! For cleavage with HpyCH4III, Bst4CI, or TaaI

35 ! cleavage is in lower chain before base 88.

! 77 788 888 888 889 999 999 9

! 78 901 234 567 890 123 456 7

T<sub>m</sub><sup>w</sup>

T<sub>m</sub><sup>K</sup>

	(H43.77.97.1-02#1)	5'-cc gtg tat tAC TGT gcg aga g-3'	6462.6
40	(H43.77.97.1-03#2)	5'-ct gtg tat tAC TGT gcg aga g-3'	6260.6
	(H43.77.97.108#3)	5'-cc gtg tat tAC TGT gcg aga g-3'	6462.6
	(H43.77.97.323#22)	5'-cc gta tat tac tgt gcg aaa g-3'	6058.7
	(H43.77.97.330#23)	5'-ct gtg tat tac tgt gcg aaa g-3'	6058.7
	(H43.77.97.439#44)	5'-ct gtg tat tac tgt gcg aga c-3'	6260.6

(H43.77.97.551#48) 5'-cc atg tat tac tgt gcg aga -3' 6260.6  
(H43.77.97.5a#49) 5'-cc atg tat tAC TGT gcg aga -3' 5858.3

## F.2 Extender and Bridges

! XbaI and AflIII sites in bridges are bunged

- 5 (H43.XABr1) 5'-gggtgtagtga-  
|TCT|AGt|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-  
|aac|agC|TTt|AGg|gct|qag|qac|aCT|GCA|Gtc|tac|tat|tgt gcg aga-3'  
(H43.XABr2) 5'-gggtgtagtga-  
|TCT|AGt|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-  
10 |aac|agC|TTt|AGg|gct|qag|qac|aCT|GCA|Gtc|tac|tat|tgt gcg aaa-3'  
(H43.XAExt) 5'-ATAgTAgAcT gcAgTgTccT cAgcccTTAA gcTgTTcATc  
TgcAAgTAgA-  
gAgTATTcTT AgAgTTgTcT cTAgATcAcT AcAcc-3'  
!H43.XAExt is the reverse complement of  
15 ! 5'-gggtgtagtga-  
! |TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-  
! |aac|agC|TTA|AGg|gct|qag|qac|aCT|GCA|Gtc|tac|tat|-3'  
  
(H43.XAPCR) 5'-gggtgtagtga |TCT|AGA|gac|aac-3'  
! XbaI and AflIII sites in bridges are bunged  
20 (H43.ABr1) 5'-gggtgtagtga-  
|aac|agC|TTt|AGg|gct|qag|qac|aCT|GCA|Gtc|tac|tat|tgt gcg aga-3'  
(H43.ABr2) 5'-gggtgtagtga-  
|aac|agC|TTt|AGg|gct|qag|qac|aCT|GCA|Gtc|tac|tat|tgt gcg aaa-3'  
(H43.AEExt) 5'-ATAgTAgAcTgcAgTgTccTcAgcccTTAAgcTgTTTcAcTAcAcc-3'  
25 !(H43.AEExt) is the reverse complement of 5'-gggtgtagtga-  
! |aac|agC|TTA|AGg|gct|qag|qac|aCT|GCA|Gtc|tac|tat|-3'  
(H43.APCR) 5'-gggtgtagtga |aac|agC|TTA|AGg|gct|q-3'

Table 5: Analysis of frequency of matching REadaptors in actual V genes

A: HpyCH4V in HC at bases 35-56

		Number of mismatches..... Number																Id	Probe		
Id	Ntot	0	1	2	3	4	5	6	7	8	9	10	Cut								
5	1	510	5	11	274	92	61	25	22	11	1	3	5	443	6-1	agttctcccTGCAgctgaactc					
	2	192	54	42	32	24	15	2	3	10	3	1	6	167	3-11	cactgtatcTGCAaatgaacag					
	3	58	19	7	17	6	5	1	0	1	0	2	0	54	3-09	ccctgtatcTGCAaatgaacag					
	4	267	42	33	9	8	8	82	43	22	8	11	1	100	5-51	ccgcctaccTGCAgtggagcag					
	5	250	111	59	41	24	7	5	1	0	0	2	0	242	3-15	cgctgtatcTGCAaatgaacag					
10	6	7	0	2	0	1	0	0	0	0	0	4	0	3	7-4.1	cggcatatcTGCAgatctgcag					
	7	7	0	2	2	0	0	2	1	0	0	0	0	4	3-73	cggcgtatcTGCAaatgaacag					
	8	26	10	4	1	3	1	2	1	3	1	0	0	19	5-a	ctgcctaccTGCAgtggagcag					
	9	21	8	2	3	1	6	1	0	0	0	0	0	20	3-49	tcgcctatcTGCAaatgaacag					
	1338	249	162	379	149	103	120	71	47	13	23	12	1052								
15	249	411	790	939	1162	1280	1316														
	1042	1233	1293	1338																	

Id	Probe	dotted probe
6-1	agttctccccTGCAgctgaactc	agttctcccc <b>TGCA</b> gctgaactc
3-11	caactgtatcTGCAaatgaacag	cac.g.at.....aa.....ag
3-09	ccctgtatcTGCAaatgaacag	ccc.g.at.....aa.....ag
5-51	ccgcctaccTGCAgtggagcag	ccgc..a.....tg..g.ag

- 139 -

3-15 cgctgtatcTGCAaatgaacag c.c.g.at.....aa.....ag  
7-4.1 cggcatatcTGCAgatctgcag c.gca.at.....a.ctg.ag  
3-73 cggcgtatcTGCAaatgaacag c.gcg.at.....aa.....ag  
5-a ctgcctaccTGCAgtggagcag ctgc..a.....tg..g.ag  
5 3-49 tcgcctatcTGCAaatgaacag tcgc..at.....aa.....ag

Seqs with the expected RE site only.....1004

(Counts only cases with 4 or fewer mismatches)

Seqs with only an unexpected site..... 0

Seqs with both expected and unexpected..... 48

10 (Counts only cases with 4 or fewer mismatches)

Seqs with no sites..... 0

B: Bp1 in HC

	Id												Ntot												0												1												2												3												4												5												6												7												8												Ncut												Name																							
	1												133												73												16												11												13												6												9												1												4												0												119												1-58												acatggaGCTGAGcagcctgag											
15	2												14												11												1												0												0												0												1												0												1												12												1-02												acatggagctgagcaggctgag																							
	3												34												17												8												2												6												1												0												0												0												0												0												1-18												acatggagctgaggagcctgag											
	4												120												50												32												16												10												9												1												1												0												2												5-51												acctgcagtggagcagcctgaa																							
	5												55												13												11												10												17												3												1												0												0												0												3-15												atctgcaaatgaacagcctgaa																							
	6												340												186												88												41												15												6												3												0												1												0												0												3303												atctgcaaatgaacagcctgag											
20	7												82												25												16												25												12												1												3												0												0												0												3-20												atctgcaaatgaacagtctgag																							



Name	Full sequence	Dot mode
1-58	acatggaGCTGAGCagcctgag	acatggaGCTGAGcagcctgag
1-02	acatggag <b>ctctgag</b> caggctgag	.....g.....
1-18	acatggagctgaggagcctgag	.....g.....
5-51	acctgcagtgagcagcctgaa	..c..c..tg.....a
3-15	atctgcaaatgaacagcctgaa	tc..c..aa...a.....a
3-30.3	atctgcaaatgaacagcctgag	tc..c..aa...a.....
3-20	atctgcaaatgaacagtctgag	tc..c..aa...a..t....
7-4.1	atctgcagatctgcagcctaaa	tc..c..a.ct.....a.a
3-66	atcttcaaatgaacagcctgag	tc.tc.aa...a.....
3-64	atcttcaaatggcagcctgag	tc.tc.aa..g.....
4-30.1	ccctgaagctgagctctgtgac	c.c..a.....tctg...c
6-1	ccctgcagctgaacctgtgac	c.c..c.....a.ctctg...c
2-70	tccttacaatgaccaacatgga	t.c.tacaa...c..a.a..ga
2-26	tccttacatgaccaacatgga	t.c.tacca...c..a.a..ga

Seqs with the expected RE site only.....	597 (counting sequences with 4 or fewer mismatches)
Seqs with only an unexpected site.....	2
Seqs with both expected and unexpected....	2
Seqs with no sites.....	686

C. HovCH4III, Bst4Cl, or TaaI in HC

2. HpyCH4III, Bst4CI, or Taal in HC

THE UNIVERSITY OF CHICAGO

- 142 -

Number of sequences..... 1617





	24	213	26	56	60	42	20	7	2	0	0	204	5a#49
<u>ccatgtattactgtgcgagaAA ..a.....AA</u>													

Group	337	471	363	218	130	58	23	11	6				
Cumulative	337	808	1171	1389	1519	1577	1600	1611	1617				

5    Seqs with the expected RE site only.....1511  
     Seqs with only an unexpected site..... 0  
     Seqs with both expected and unexpected.... 8  
     Seqs with no sites..... 0

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

Table 5D:

Analysis repeated using only 8 best REadaptors

	Id	Ntot	0	1	2	3	4	5	6	7	8+						
5	1	301	78	101	54	32	16	9	10	1	0	281 102#1					
		ccgtgtattactgtgcgagaga															
	2	493	69	155	125	73	37	14	11	3	6	459 103#2					
		ctgtgtattactgtgcgagaga															
	3	189	52	45	38	23	18	5	4	1	3	176 108#3					
10		ccgtgtattactgtgcgagagg															
	4	127	29	23	28	24	10	6	5	2	0	114 323#22					
		ccgtatattactgtgcgaaaga															
	5	78	21	25	14	11	1	4	2	0	0	72 330#23					
		ctgtgtattactgtgcgaaaga															
						6	79	15	17	25	8	11	1	2	0	0	76
15		439#44 ctgtgtattactgtgcgagaca															
	7	43	14	15	5	5	3	0	1	0	0	42 551#48					
		ccatgtattactgtgcgagaca															
	8	307	26	63	72	51	38	24	14	13	6	250 5a#49					
		ccatgtattactgtgcgaga															
20	1	102#1	ccgtgtattactgtgcgagaga					ccgtgtattactgtgcgagaga									
	2	103#2	ctgtgtattactgtgcgagaga					.t.....									
	3	108#3	ccgtgtattactgtgcgagagg					.....g									
	4	323#22	ccgtatattactgtgcgaaaga					....a.....a...									
	5	330#23	ctgtgtattactgtgcgaaaga					.t.....a...									
25	6	439#44	ctgtgtattactgtgcgagaca					.t.....c.									
	7	551#48	ccatgtattactgtgcgagaca					..a.....c.									
	8	5a#49	ccatgtattactgtgcgagaAA					..a.....AA									

Seqs with the expected RE site only.....1463 / 1617

Seqs with only an unexpected site..... 0

30 Seqs with both expected and unexpected.... 7

Seqs with no sites..... 0

Table 6: Human HC GLG FR1 Sequences

VH Exon - Nucleotide sequence alignment

VH1	
5	1-02 CAG GTG CAG CTG GTG CAG TCT GGG GCT GAG GTG AAG AAG CCT GGG GCC TCA GTG AAG GTC TCC TGC AAG GCT TCT GGA TAC ACC TTC ACC
	1-03 cag gtC cag ctT gtg cag tct ggg gct gag gtg aag aag cct ggg gcc tca gtg aag gtT tcc tgc aag gct tct gga tac acc ttc acT
	1-08 cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg gcc tca gtg aag gtc tcc tgc aag gct tct gga tac acc ttc acc
10	1-18 cag gtT cag ctg gtg cag tct ggA gct gag gtg aag aag cct ggg gcc tca gtg aag gtc tcc tgc aag gct tct ggT tac acc ttT acc
	1-24 cag gtC cag ctg gtA cag tct ggg gct gag gtg aag aag cct ggg gcc tca gtg aag gtc tcc tgc aag gTt tC gga tac acc CtC acT
	1-45 cag Atg cag ctg gtg cag tct ggg gct gag gtg aag aag Act ggg Tcc tca gtg aag gtT tcc tgc aag gct tC gga tac acc ttc acc
15	1-46 cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg gcc tca gtg aag gtT tcc tgc aag gCt tct gga tac acc ttc acc
	1-58 caA Atg cag ctg gtg cag tct ggg Cct gag gtg aag aag cct ggg Acc tca gtg aag gtc tcc tgc aag gct tct gga tTc acc ttT acT
	20 1-69 cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg Tcc tCg gtg aag gtc tcc tgc aag gct tct gga GGc acc ttc aGc
25	1-e cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg Tcc tCg gtg aag gtc tcc tgc aag gct tct gga GGc acc ttc aGc
	1-f Gag gtC cag ctg gtA cag tct ggg gct gag gtg aag aag cct ggg gCt Aca gtg aaA
	Atc tcc tgc aag gTt tct gga tac acc ttc acc
VH2	
30	2-05 CAG ATC ACC TTG AAG GAG TCT GGT CCT ACG CTG GTG AAA CCC ACA CAG ACC CTC ACG CTG ACC TGC ACC TTC TCT GGG TTC TCA CTC AGC
	2-26 cag Gtc acc ttg aag gag tct ggt cct GTg ctg gtg aaa ccc aca Gag acc ctc acg ctg acc tgc acc Gtc tct ggg ttc tca ctc agc
	2-70 cag Gtc acc ttg aag gag tct ggt cct Gcg ctg gtg aaa ccc aca cag acc ctc acA ctg acc tgc acc ttc tct ggg ttc tca ctc agc
VH3	
35	3-07 GAG GTG CAG CTG GTG GAG TCT GGG GGA GGC TTG GTC CAG CCT GGG GGG TCC CTG AGA CTC TCC TGT GCA GCC TCT GGA TTC ACC TTT AGT
	3-09 gaA gtg cag ctg gtg gag tct ggg gga ggc ttg gtA cag cct ggC Agg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttt GAt
	3-11 Cag gtg cag ctg gtg gag tct ggg gga ggc ttg gtc Aag cct ggA ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt
40	3-13 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtA cag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt
	3-15 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtA Aag cct ggg ggg tcc ctT aga ctc tcc tgt gca gcc tct gga ttc acT ttC agt



4-30.4 cag gtg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcA CAg acc ctg tcc  
ctc acc tgc Act gtc tct ggt ggc tcc atc agc

4-31 cag gtg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcA CAg acc ctg tcc  
ctc acc tgc Act gtc tct ggt ggc tcc atc agc

5 4-34 cag gtg cag ctA cag Cag tGg ggc Gca gga ctg Ttg aag cct tcg gAg acc ctg tcc  
ctc acc tgc gct gtc tAt ggt ggG tcc Ttc agT

4-39 cag Ctg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcg gAg acc ctg tcc  
ctc acc tgc Act gtc tct ggt ggc tcc atc agc

4-59 cag gtg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcg gAg acc ctg tcc  
ctc acc tgc Act gtc tct ggt ggc tcc atc agT

10 4-61 cag gtg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcg gAg acc ctg tcc  
ctc acc tgc Act gtc tct ggt ggc tcc Gtc agc

4-b cag gtg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcg gAg acc ctg tcc  
ctc acc tgc gct gtc tct ggt TAc tcc atc agc

15 VH5

5-51 GAG GTG CAG CTG GTG CAG TCT GGA GCA GAG GTG AAA AAG CCC GGG GAG TCT CTG AAG  
ATC TCC TGT AAG GGT TCT GGA TAC AGC TTT ACC

5-a gaA gtg cag ctg gtg cag tct gga gca gag gtg aaa aag ccc ggg gag tct ctg aGg  
atc tcc tgt aag ggt tct gga tac agc ttt acc

20 VH6

6-1 CAG GTA CAG CTG CAG CAG TCA GGT CCA GGA CTG GTG AAG CCC TCG CAG ACC CTC TCA  
CTC ACC TGT GCC ATC TCC GGG GAC AGT GTC TCT

VH7

7-4.1 CAG GTG CAG CTG GTG CAA TCT GGG TCT GAG TTG AAG AAG CCT GGG GCC TCA GTG AAG  
GTT TCC TGC AAG GCT TCT GGA TAC ACC TTC ACT

25

Table 7: RERS sites in Human HC GLG FR1s where there are at least 20 GLGs cut

BsgI GTGCAG		71 (cuts 16/14 bases to right)				
5	1: 4	1: 13	2: 13	3: 4	3: 13	4: 13
	6: 13	7: 4	7: 13	8: 13	9: 4	9: 13
	10: 4	10: 13	15: 4	15: 65	16: 4	16: 65
	17: 4	17: 65	18: 4	18: 65	19: 4	19: 65
	20: 4	20: 65	21: 4	21: 65	22: 4	22: 65
10	23: 4	23: 65	24: 4	24: 65	25: 4	25: 65
	26: 4	26: 65	27: 4	27: 65	28: 4	28: 65
	29: 4	30: 4	30: 65	31: 4	31: 65	32: 4
	32: 65	33: 4	33: 65	34: 4	34: 65	35: 4
	35: 65	36: 4	36: 65	37: 4	38: 4	39: 4
15	41: 4	42: 4	43: 4	45: 4	46: 4	47: 4
	48: 4	48: 13	49: 4	49: 13	51: 4	
	There are 39 hits at base# 4					
	There are 21 hits at base# 65					
20	-"- ctgcac			9		
	12: 63	13: 63	14: 63	39: 63	41: 63	42: 63
	44: 63	45: 63	46: 63			
	BbvI GCAGC			65		
	1: 6	3: 6	6: 6	7: 6	8: 6	9: 6
25	10: 6	15: 6	15: 67	16: 6	16: 67	17: 6
	17: 67	18: 6	18: 67	19: 6	19: 67	20: 6
	20: 67	21: 6	21: 67	22: 6	22: 67	23: 6
	23: 67	24: 6	24: 67	25: 6	25: 67	26: 6
	26: 67	27: 6	27: 67	28: 6	28: 67	29: 6
30	30: 6	30: 67	31: 6	31: 67	32: 6	32: 67
	33: 6	33: 67	34: 6	34: 67	35: 6	35: 67
	36: 6	36: 67	37: 6	38: 6	39: 6	40: 6
	41: 6	42: 6	43: 6	44: 6	45: 6	46: 6
	47: 6	48: 6	49: 6	50: 12	51: 6	
There are 43 hits at base# 6 Boded sites very near sites						
listed below						
35	There are 21 hits at base# 67					
	-"- gctgc			13		
	37: 9	38: 9	39: 9	40: 3	40: 9	41: 9
	42: 9	44: 3	44: 9	45: 9	46: 9	47: 9

Table 7: RERS sites in Human HC GLG FR1s where there are at least 20 GLGs cut





There are 2 hits at base# 3

There are 1 hits at base# 12

There are 21 hits at base# 67

5 MspAII CMGckg 48

1:	7	3:	7	4:	7	5:	7	6:	7	7:	7
8:	7	9:	7	10:	7	11:	7	15:	7	16:	7
17:	7	18:	7	19:	7	20:	7	21:	7	22:	7
23:	7	24:	7	25:	7	26:	7	27:	7	28:	7
29:	7	30:	7	31:	7	32:	7	33:	7	34:	7
35:	7	36:	7	37:	7	38:	7	39:	7	<u>40:</u>	<u>1</u>
<u>40:</u>	<u>7</u>	41:	7	42:	7	<u>44:</u>	<u>1</u>	<u>44:</u>	<u>7</u>	45:	7
46:	7	47:	7	48:	7	49:	7	50:	7	51:	7

There are 46 hits at base# 7

15 PvuII CAGctg 48

1:	7	3:	7	4:	7	5:	7	6:	7	7:	7
8:	7	9:	7	10:	7	11:	7	15:	7	16:	7
17:	7	18:	7	19:	7	20:	7	21:	7	22:	7
23:	7	24:	7	25:	7	26:	7	27:	7	28:	7
29:	7	30:	7	31:	7	32:	7	33:	7	34:	7
35:	7	36:	7	37:	7	38:	7	39:	7	<u>40:</u>	<u>1</u>
<u>40:</u>	<u>7</u>	41:	7	42:	7	<u>44:</u>	<u>1</u>	<u>44:</u>	<u>7</u>	45:	7
46:	7	47:	7	48:	7	49:	7	50:	7	51:	7

25 There are 46 hits at base# 7  
There are 2 hits at base# 1

AluI AGct 54

1:	8	2:	8	3:	8	4:	8	4:	24	5:	8
6:	8	7:	8	8:	8	9:	8	10:	8	11:	8
15:	8	16:	8	17:	8	18:	8	19:	8	20:	8
21:	8	22:	8	23:	8	24:	8	25:	8	26:	8
27:	8	28:	8	29:	8	29:	69	30:	8	31:	8
32:	8	33:	8	34:	8	35:	8	36:	8	37:	8
38:	8	39:	8	<u>40:</u>	<u>2</u>	<u>40:</u>	<u>8</u>	41:	8	42:	8
43:	8	<u>44:</u>	<u>2</u>	<u>44:</u>	<u>8</u>	45:	8	46:	8	47:	8
48:	8	48:	82	49:	8	49:	82	50:	8	51:	8



There are 35 hits at base# 39 39 and 40 together twice.

There are 2 hits at base# 40

BsaJI Ccnnngg 47

5	1: 40	2: 40	3: 40	4: 40	5: 40	7: 40
	8: 40	9: 40	9: 47	10: 40	10: 47	11: 40
	15: 40	18: 40	19: 40	20: 40	21: 40	22: 40
	23: 40	24: 40	25: 40	26: 40	27: 40	28: 40
	29: 40	30: 40	31: 40	32: 40	34: 40	35: 20
10	35: 40	36: 40	37: 24	38: 24	39: 24	41: 24
	42: 24	44: 24	45: 24	46: 24	47: 24	<u>48: 40</u>
	<u>48: 41</u>	<u>49: 40</u>	<u>49: 41</u>	50: 74	51: 40	

There are 32 hits at base# 40 40 and 41 together twice

There are 2 hits at base# 41

15 There are 9 hits at base# 24

There are 2 hits at base# 47

BstNI CCwgg 44

PspGI ccwgg

20 ScrFI(\$M.HpaII) CCwgg

	1: 40	2: 40	3: 40	4: 40	5: 40	7: 40
	8: 40	9: 40	10: 40	11: 40	15: 40	16: 40
	17: 40	18: 40	19: 40	20: 40	21: 30	21: 40
	22: 40	23: 40	24: 40	25: 40	26: 40	27: 40
25	28: 40	29: 40	30: 40	31: 40	32: 40	33: 40
	34: 40	35: 40	36: 40	37: 25	38: 25	39: 25
	41: 25	42: 25	44: 25	45: 25	46: 25	47: 25
	50: 25	51: 40				

There are 33 hits at base# 40

30

ScrFI CCnngg 50

	1: 40	2: 40	3: 40	4: 40	5: 40	7: 40
	8: 40	9: 40	10: 40	11: 40	15: 40	16: 40
	17: 40	18: 40	19: 40	20: 40	21: 30	21: 40
35	22: 40	23: 40	24: 40	25: 40	26: 40	27: 40
	28: 40	29: 40	30: 40	31: 40	32: 40	33: 40
	34: 40	35: 20	35: 40	36: 40	37: 25	38: 25
	39: 25	41: 25	42: 25	44: 25	45: 25	46: 25

There are 2 hits at base# 41

5

10

There are 11 hits at base# 43

15

20

There are 17 hits at base# 46      There are 11 hits at base# 43

30

35

41: 22    42: 21    42: 22    43: 80    44: 21    44: 22    45: 21    45: 22





26: 77 27: 77 28: 77 29: 77 31: 77 32: 77  
33: 77 35: 77 36: 77

There are 21 hits at base# 77

5 MlyI GAGTC 38  
12: 16 13: 16 14: 16 15: 16 16: 16 17: 16  
18: 16 19: 16 20: 16 21: 16 22: 16 23: 16  
24: 16 25: 16 26: 16 27: 16 27: 26 28: 16  
29: 16 31: 16 32: 16 33: 16 34: 16 35: 16  
10 36: 16 36: 26 37: 16 38: 16 39: 16 40: 16  
41: 16 42: 16 44: 16 45: 16 46: 16 47: 16  
48: 46 49: 46

There are 34 hits at base# 16

15 -"- GACTC 21  
15: 56 16: 56 17: 56 18: 56 19: 56 20: 56  
21: 56 22: 56 23: 56 24: 56 25: 56 26: 56  
27: 56 28: 56 29: 56 30: 56 31: 56 32: 56  
33: 56 35: 56 36: 56

20 There are 21 hits at base# 56

PleI gagtc 38  
12: 16 13: 16 14: 16 15: 16 16: 16 17: 16  
18: 16 19: 16 20: 16 21: 16 22: 16 23: 16  
25 24: 16 25: 16 26: 16 27: 16 27: 26 28: 16  
29: 16 31: 16 32: 16 33: 16 34: 16 35: 16  
36: 16 36: 26 37: 16 38: 16 39: 16 40: 16  
41: 16 42: 16 44: 16 45: 16 46: 16 47: 16  
48: 46 49: 46

30 There are 34 hits at base# 16

-"- gactc 21  
15: 56 16: 56 17: 56 18: 56 19: 56 20: 56  
21: 56 22: 56 23: 56 24: 56 25: 56 26: 56  
27: 56 28: 56 29: 56 30: 56 31: 56 32: 56  
35 33: 56 35: 56 36: 56

There are 21 hits at base# 56

AlwNI CAGNNNctg 26  
15: 68 16: 68 17: 68 18: 68 19: 68 20: 68

- 158 -

21: 68	22: 68	23: 68	24: 68	25: 68	26: 68
27: 68	28: 68	29: 68	30: 68	31: 68	32: 68
33: 68	34: 68	35: 68	36: 68	39: 46	40: 46
41: 46	42: 46				

5    **There are 22 hits at base# 68**

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100



Table 8: Kappa FR1 GLGs

5

10

1520253035

	GTC ATC TGG ATG ACC CAG TCT CCA TCC TTA CTC TCT	
	GCA TCT ACA GGA GAC AGA GTC ACC ATC AGT TGT !	L24
	GCC ATC CAG ATG ACC CAG TCT CCA TCC TCC CTG TCT	
	GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGC !	L11
5	GAC ATC CAG ATG ACC CAG TCT CCT TCC ACC CTG TCT	
	GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGC !	L12
	GAT ATT GTG ATG ACC CAG ACT CCA CTC TCC CTG CCC	
	GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC !	O11
	GAT ATT GTG ATG ACC CAG ACT CCA CTC TCC CTG CCC	
10	GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC !	O1
	GAT GTT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC	
	GTC ACC CTT GGA CAG CCG GCC TCC ATC TCC TGC !	A17
	GAT GTT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC	
	GTC ACC CTT GGA CAG CCG GCC TCC ATC TCC TGC !	A1
15	GAT ATT GTG ATG ACC CAG ACT CCA CTC TCT CTG TCC	
	GTC ACC CCT GGA CAG CCG GCC TCC ATC TCC TGC !	A18
	GAT ATT GTG ATG ACC CAG ACT CCA CTC TCT CTG TCC	
	GTC ACC CCT GGA CAG CCG GCC TCC ATC TCC TGC !	A2
	GAT ATT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC	
20	GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC !	A19
	GAT ATT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC	
	GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC !	A3
	GAT ATT GTG ATG ACC CAG ACT CCA CTC TCC TCA CCT	
	GTC ACC CTT GGA CAG CCG GCC TCC ATC TCC TGC !	A23
25	GAA ATT GTG TTG ACG CAG TCT CCA GGC ACC CTG TCT	
	TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC !	A27
	GAA ATT GTG TTG ACG CAG TCT CCA GCC ACC CTG TCT	
	TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC !	A11
	GAA ATA GTG ATG ACG CAG TCT CCA GCC ACC CTG TCT	
30	GTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC !	L2
	GAA ATA GTG ATG ACG CAG TCT CCA GCC ACC CTG TCT	
	GTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC !	L16
	GAA ATT GTG TTG ACA CAG TCT CCA GCC ACC CTG TCT	
	TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC !	L6
35	GAA ATT GTG TTG ACA CAG TCT CCA GCC ACC CTG TCT	



Table 9 RERS sites found in Human Kappa FRI GLGs

VKI	MslI	FokI --> <-- -->	PstFI	BsrI	BsmAI	MnlI	HpyCH 4V
O12 1-69	3	3 23	12 49	15	18 47	26	36
O2 101-169	103	103 123	112 149	115	118 147	126	136
O18 201-269	203	203 223	212 249	215	218 247	226	236
O8 301-369	303	303 323	312 349	315	318 347	326	336
A20 401-469	403	403 423	412 449	415	418 447	426	436
A30 501-569	503	503 523	512 549	515	518 547	526	536
L14 601-669	603	603	612 649	615	618 647	-	636
L1 701-769	703	703 723	712 749	715	718 747	726	736
L15 801-869	803	803 823	812 849	815	818 847	826	836
L4 901-969	-	903 923	912 949	906 915	918 947	926	936
L18 1001-1069	-	1003	1012 1049	1006 1015	1018 1047	1026	1036
L5 1101-1169	1103	-	1112 1149	1115	1118 1147	-	1136
L19 1201-1269	1203	1203	1212 1249	1215	1218 1247	-	1236
L8 1301-1369	-	1303 1323	1312 1349	1306 1315	1318 1347	-	1336
L23 1401-1469	1403	1403 1408	1412 1449	1415	1418 1447	-	1436
L9 1501-1569	1503	1503 1508 1523	1512 1549	1515	1518 1547	1526	1536

5

10

15

	MslI	FokI --> <-- -->	PstI	BsrI	BsmAI	MnlI	HpyCH 4V
L24 1601-1669	1603	1608 1623	1612 1649	1615	1618 1647	-	1636
L11 1701-1769	1703	1703 1723	1712 1749	1715	1718 1747	1726	1736
L12 1801-1869	1803	1803	1812 1849	1815	1818 1847	-	1836
VKII							
O11 1901-1969	-	-	-	-	-	1956	-
O1 2001-2069	-	-	-	-	-	2056	-
A17 2101-2169	-	-	2112	-	2118	2156	-
A1 2201-2269	-	-	2212	-	2218	2256	-
A18 2301-2369	-	-	-	-	-	2356	-
A2 2401-2469	-	-	-	-	-	2456	-
A19 2501-2569	-	-	2512	-	2518	2556	-
A3 2601-2669	-	-	2612	-	2618	2656	-
A23 2701-2769	-	-	-	-	-	2729 2756	-
VKIII							
A27 2801-2869	-	-	2812	-	2818 2839	2860	-
A11 2901-2969	-	-	2912	-	2918 2939	2960	-
L2 3001-3069	-	-	3012	-	3018 3039	3060	-
L16 3101-3169	-	-	3112	-	3118 3139	3160	-



Table 9 RERS sites found in Human Kappa FR1 GLGs, continued

	SfaNI	SfcI	HinfI	MlyI --> --> <--	MaeIII Tsp45I same sites	HphI xx38 xx56 xx62	HpaII MspI xx06 xx52
VKI							
O12 1-69	37	41	53	53	55	56	-
O2 101-169	137	141	153	153	155	156	-
O18 201-269	237	241	253	253	255	256	-
O8 301-369	337	341	353	353	355	356	-
A20 401-469	437	441	453	453	455	456	-
A30 501-569	537	541	553	553	555	556	-
L14 601-669	637	641	653	653	655	656	-
L1 701-769	737	741	753	753	755	756	-
L15 801-869	837	841	853	853	855	856	-
L4 901-969	937	941	953	953	955	956	-
L18 1001-1069	1037	1041	1053	1053	1055	1056	-
L5 1101-1169	1137	1141	1153	1153	1155	1156	-
L19 1201-1269	1237	1241	1253	1253	1255	1256	-
L8 1301-1369	1337	1341	1353	1353	1355	1356	-
L23 1401-1469	1437	1441	1453	1453	1455	1456	1406





	SfaNI	SfcI	HinfI	MlyI --> --> <--	MaeIII Tsp45I same sites	HphI xx38 xx56 xx62	HpaII MspI xx06 xx52
L2 3001-3069	-	-	-	-			-
L16 3101-3169	-	-	-	-			-
L6 3201-3269	-	-	-	-			-
L20 3301-3369	-	-	-	-			-
L25 3401-3469	-	-	-	-			-
VKTIV							
B3 3501-3569	-	-	3525	3525			-
VKV							
B2 3601-3669	-	-	3639	3639			-
VKVI							
A26 3701-3769	-	-	3712 3739	3712 3739	3737 3755	3756 3762	-
A10 3801-3869	-	-	3812 3839	3812 3839	3837 3855	3856 3862	-
A14 3901-3969	-	-	3939	3939	3937 3955	3956 3962	-

Table 9 RERS sites found in Human Kappa FR1, continued

	BsaJI xx29 xx42 xx43	BssKI (NstNI) xx22 xx30 xx43	BpmI xx20 xx41 xx44 --> --> <--	BstFI Cac8I NaeI NgoMTV	HaeIII	Tsp509I
YKI	-	-	-	-	-	-
O12 1-69	-	-	-	-	-	-
O2 101-169	-	-	-	-	-	-
O18 201-269	-	-	-	-	-	-
O8 301-369	-	-	-	-	-	-
A20 401-469	-	-	-	-	-	-
A30 501-569	-	-	-	-	-	-
L14 601-669	-	-	-	-	-	-
L1 701-769	-	-	-	-	-	-
L15 801-869	-	-	-	-	-	-
L4 901-969	-	-	-	-	-	-
L18 1001-1069	-	-	-	-	-	-
L5 1101-1169	-	-	-	-	-	-
L19 1201-1269	-	-	-	-	-	-
L8 1301-1369	-	-	-	-	-	-

5

10

15

	BsaJI xx29 xx42 xx43	BssKI (NstNII) xx22 xx30 xx43	BpmI xx20 xx41 xx44 --> --> <--	BstFI Cac8I NaeI NgoMIV	HaeIII	Tsp509I
L23 1401-1469	-	-	-	-	-	-
L9 1501-1569	-	-	-	-	-	-
L24 1601-1669	-	-	-	-	-	-
L11 1701-1769	-	-	-	-	-	-
L12 1801-1869	-	-	-	-	-	-
VKII						
O11 1901-1969	1942	1943	1944	1951	1954	-
O1 2001-2069	2042	2043	2044	2051	2054	-
A17 2101-2169	2142	-	-	2151	2154	-
A1 2201-2269	2242	-	-	2251	2254	-
A18 2301-2369	2342	2343	-	2351	2354	-
A2 2401-2469	2442	2443	-	2451	2454	-
A19 2501-2569	2542	2543	2544	2551	2554	-
A3 2601-2669	2642	2643	2644	2651	2654	-
A23 2701-2769	2742	-	-	2751	2754	-
VKIII						

5

10

15



Table 10 Lambda FR1 GLG sequences

! VL1

5 CAG TCT GTG CTG ACT CAG CCA CCC TCG GTG TCT GAA  
GCC CCC AGG CAG AGG GTC ACC ATC TCC TGT ! 1a  
cag tct gtg ctg acG cag ccG ccc tcA gtg tct gGG  
gcc ccA Ggg cag agg gtc acc atc tcc tgC ! 1e  
cag tct gtg ctg act cag cca ccc tcA gCg tct gGG  
Acc ccc Ggg cag agg gtc acc atc tcT tgt ! 1c  
cag tct gtg ctg act cag cca ccc tcA gCg tct gGG  
10 Acc ccc Ggg cag agg gtc acc atc tcT tgt ! 1g  
cag tct gtg Ttg acG cag ccG ccc tcA gtg tct gCG  
gcc ccA GgA cag aAg gtc acc atc tcc tgC ! 1b

! VL2

15 CAG TCT GCC CTG ACT CAG CCT CCC TCC GCG TCC GGG  
TCT CCT GGA CAG TCA GTC ACC ATC TCC TGC ! 2c  
cag tct gcc ctg act cag cct cGc tcA gTg tcc ggg  
tct cct gga cag tca gtc acc atc tcc tgc! 2e  
cag tct gcc ctg act cag cct Gcc tcc gTg tcT ggg  
tct cct gga cag tcG Atc acc atc tcc tgc ! 2a2  
20 cag tct gcc ctg act cag cct ccc tcc gTg tcc ggg  
tct cct gga cag tca gtc acc atc tcc tgc ! 2d  
cag tct gcc ctg act cag cct Gcc tcc gTg tcT ggg  
tct cct gga cag tcG Atc acc atc tcc tgc ! 2b2

! VL3

25 TCC TAT GAG CTG ACT CAG CCA CCC TCA GTG TCC GTG  
TCC CCA GGA CAG ACA GCC AGC ATC ACC TGC! 3r  
tcc tat gag ctg act cag cca cTc tca gtg tcA gtg  
Gcc cTG gga cag acG gcc agG atT acc tgT ! 3j  
tcc tat gag ctg acA cag cca ccc tcG gtg tcA gtg  
30 tcc cca gga caA acG gcc agG atc acc tgc! 3p  
tcc tat gag ctg acA cag cca ccc tcG gtg tcA gtg  
tcc cTa gga cag aTG gcc agG atc acc tgc ! 3a  
tcT tCt gag ctg act cag GAC ccT GcT gtg tcT gtg  
Gcc TTG gga cag aca gTc agG atc acA tgc ! 3l

tcc tat gTg ctg act cag cca ccc tca gtg tcA gtg  
Gcc cca gga Aag acG gcc agG atT acc tgT ! 3h  
tcc tat gag ctg acA cag cTa ccc tcG gtg tcA gtg  
tcc cca gga cag aca gcc agG atc acc tgc ! 3e  
5 tcc tat gag ctg aTG cag cca ccc tcG gtg tcA gtg  
tcc cca gga cag acG gcc agG atc acc tgc ! 3m  
tcc tat gag ctg acA cag cca Tcc tca gtg tcA gtg  
tcT ccG gga cag aca gcc agG atc acc tgc ! V2-19  
! VL4  
10 CTG CCT GTG CTG ACT CAG CCC CCG TCT GCA TCT GCC  
TTG CTG GGA GCC TCG ATC AAG CTC ACC TGC ! 4c  
cAg cct gtg ctg act caA TcA TcC tct gcC tct gcT  
tCC ctg gga Tcc tcg Gtc aag ctc acc tgc ! 4a  
cAg cTt gtg ctg act caA TcG ccC tct gcC tct gcc  
15 tCC ctg gga gcc tcg Gtc aag ctc acc tgc ! 4b  
! VL5  
CAG CCT GTG CTG ACT CAG CCA CCT TCC TCC TCC GCA  
TCT CCT GGA GAA TCC GCC AGA CTC ACC TGC ! 5e  
cag Gct gtg ctg act cag ccG Gct tcc CTc tcT gca  
20 tct cct gga gCa tcA gcc agT ctc acc tgc ! 5c  
cag cct gtg ctg act cag cca Tct tcc CAT tcT gca  
tct Tct gga gCa tcA gTc aga ctc acc tgc ! 5b  
! VL6  
AAT TTT ATG CTG ACT CAG CCC CAC TCT GTG TCG GAG  
25 TCT CCG GGG AAG ACG GTA ACC ATC TCC TGC ! 6a  
! VL7  
CAG ACT GTG GTG ACT CAG GAG CCC TCA CTG ACT GTG  
TCC CCA GGA GGG ACA GTC ACT CTC ACC TGT ! 7a  
cag Gct gtg gtg act cag gag ccc tca ctg act gtg  
30 tcc cca gga ggg aca gtc act ctc acc tgt ! 7b  
! VL8  
CAG ACT GTG GTG ACC CAG GAG CCA TCG TTC TCA GTG  
TCC CCT GGA GGG ACA GTC ACA CTC ACT TGT ! 8a



Table 11 RERSs found in human lambda FR1 GLGs

! There are 31 lambda GLGs

	MlyI	NnnnnnGACTC	25				
	1:	6	3:	6	4:	6	6: 6 7: 6 8: 6
5	9:	6	10:	6	11:	6	12: 6 15: 6 16: 6
	20:	6	21:	6	22:	6	23: 6 23: 50 24: 6
	25:	6	25: 50	26:	6	27:	6 28: 6 30: 6
	31:	6					
	There are 23 hits at base# 6						
10	---	GAGTCNNNNNn	1				
	26:	34					
	MwoI	GCNNNNNNngc	20				
15	1:	9	2:	9	3:	9	4: 9 11: 9 11: 56
	12:	9	13:	9	14:	9	16: 9 17: 9 18: 9
	19:	9	20:	9	23:	9	24: 9 25: 9 26: 9
	30:	9	31:	9			
	There are 19 hits at base# 9						
20	HinfI	Gantc	27				
	1:	12	3:	12	4:	12	6: 12 7: 12 8: 12
	9:	12	10:	12	11:	12	12: 12 15: 12 16: 12
	20:	12	21:	12	22:	12	23: 12 23: 46 23: 56
	24:	12	25:	12	25: 56	26:	12 26: 34 27: 12
25	28:	12	30:	12	31:	12	
	There are 23 hits at base# 12						
	PleI	gactc	25				
	1:	12	3:	12	4:	12	6: 12 7: 12 8: 12
	9:	12	10:	12	11:	12	12: 12 15: 12 16: 12
30	20:	12	21:	12	22:	12	23: 12 23: 56 24: 12
	25:	12	25: 56	26:	12	27:	12 28: 12 30: 12
	31:	12					
	There are 23 hits at base# 12						
35	---	gagtc	1				



26: 34

DdeI Ctnag 32

1: 14	2: 24	3: 14	3: 24	4: 14	4: 24
5: 24	6: 14	7: 14	7: 24	8: 14	9: 14
10: 14	11: 14	11: 24	12: 14	12: 24	15: 5
15: 14	16: 14	16: 24	19: 24	20: 14	23: 14
24: 14	25: 14	26: 14	27: 14	28: 14	29: 30
30: 14	31: 14				

10 There are 21 hits at base# 14

BsaJI Ccnngg 38

1: 23	1: 40	2: 39	2: 40	3: 39	3: 40
4: 39	4: 40	5: 39	11: 39	12: 38	12: 39
13: 23	13: 39	14: 23	14: 39	15: 38	16: 39
17: 23	17: 39	18: 23	18: 39	21: 38	21: 39
21: 47	22: 38	22: 39	22: 47	26: 40	27: 39
28: 39	29: 14	29: 39	30: 38	30: 39	30: 47
31: 23	31: 32				

20 There are 17 hits at base# 39

There are 5 hits at base# 38

There are 5 hits at base# 40 Makes cleavage ragged.

MnlI cctc 35

1: 23	2: 23	3: 23	4: 23	5: 23	6: 19
6: 23	7: 19	8: 23	9: 19	9: 23	10: 23
11: 23	13: 23	14: 23	16: 23	17: 23	18: 23
19: 23	20: 47	21: 23	21: 29	21: 47	22: 23
22: 29	22: 35	22: 47	23: 26	23: 29	24: 27
27: 23	28: 23	30: 35	30: 47	31: 23	

30 There are 21 hits at base# 23

There are 3 hits at base# 19

There are 3 hits at base# 29

There are 1 hits at base# 26

There are 1 hits at base# 27 These could make cleavage ragged.

35 -- gagg



There are 4 hits at base# 39

There are 1 hits at base# 41

ScrFI CCngg

39

5 1: 41 2: 40 3: 40 3: 41 4: 40 4: 41  
 5: 40 6: 32 6: 40 7: 32 7: 40 8: 40  
 9: 32 9: 40 10: 40 11: 40 12: 39 12: 53  
 13: 40 13: 53 14: 53 16: 40 16: 53 17: 40  
 17: 53 18: 40 18: 53 19: 40 19: 53 21: 39  
 10 22: 39 23: 40 24: 40 26: 40 27: 40 28: 40  
 29: 15 29: 40 30: 39

There are 21 hits at base# 40

There are 4 hits at base# 39

There are 3 hits at base# 41

15

MaeIII gtnac

16

1: 52 2: 52 3: 52 4: 52 5: 52 6: 52  
 7: 52 9: 52 26: 52 27: 10 27: 52 28: 10  
 28: 52 29: 10 29: 52 30: 52

20 There are 13 hits at base# 52

Tsp45I gtsac

15

1: 52 2: 52 3: 52 4: 52 5: 52 6: 52  
 7: 52 9: 52 27: 10 27: 52 28: 10 28: 52  
 25 29: 10 29: 52 30: 52

There are 12 hits at base# 52

HphI tcacc

26

1: 53 2: 53 3: 53 4: 53 5: 53 6: 53  
 30 7: 53 8: 53 9: 53 10: 53 11: 59 13: 59  
 14: 59 17: 59 18: 59 19: 59 20: 59 21: 59  
 22: 59 23: 59 24: 59 25: 59 27: 59 28: 59  
 30: 59 31: 59

There are 16 hits at base# 59



Table 12: Matches to URE FR3 adapters in 79 human HC.

A. List of Heavy-chains genes sampled

	AF008566	AF103367	HSA235674	HSU94417	S83240
	AF035043	AF103368	HSA235673	HSU94418	SABVH369
5	AF103026	AF103369	HSA240559	HSU96389	SADEIGVH
	af103033	AF103370	HSCB201	HSU96391	SAH2IGVH
	AF103061	af103371	HSIGGVHC	HSU96392	SDA3IGVH
	Af103072	AF103372	HSU44791	HSU96395	SIGVHTTD
	af103078	AF158381	HSU44793	HSZ93849	SUK4IGVH
10	AF103099	E05213	HSU82771	HSZ93850	
	AF103102	E05886	HSU82949	HSZ93851	
	AF103103	E05887	HSU82950	HSZ93853	
	AF103174	HSA235661	HSU82952	HSZ93855	
	AF103186	HSA235664	HSU82961	HSZ93857	
15	af103187	HSA235660	HSU86522	HSZ93860	
	AF103195	HSA235659	HSU86523	HSZ93863	
	af103277	HSA235678	HSU92452	MCOMFRAA	
	af103286	HSA235677	HSU94412	MCOMFRVA	
	AF103309	HSA235676	HSU94415	S82745	
20	af103343	HSA235675	HSU94416	S82764	

Table 12B. Testing all distinct GLGs from bases 89.1 to 93.2 of the heavy variable domain

	Id	Nb	0	1	2	3	4		SEQ ID
NO:									
25	1	38	15	11	10	0	2	Seq1	gtgtattactgtgc 25
	2	19	7	6	4	2	0	Seq2	gtAtattactgtgc 26
	3	1	0	0	1	0	0	Seq3	gtgtattactgtAA 27
	4	7	1	5	1	0	0	Seq4	gtgtattactgtAc 28
	5	0	0	0	0	0	0	Seq5	Ttgtattactgtgc 29
30	6	0	0	0	0	0	0	Seq6	TtgtatCactgtgc 30
	7	3	1	0	1	1	0	Seq7	ACAtattactgtgc 31
	8	2	0	2	0	0	0	Seq8	ACgtattactgtgc 32
	9	9	2	2	4	1	0	Seq9	ATgtattactgtgc 33
	Group		26	26	21	4	2		
35	Cumulative		26	52	73	77	79		

Table 12C Most important URE recognition seqs in FR3 Heavy

	1	VHSzy1	GTGtattactgtgc	(ON_SHC103)	(SEQ ID NO:25)
	2	VHSzy2	GTAtattactgtgc	(ON_SHC323)	(SEQ ID NO:26)
	3	VHSzy4	GTGtattactgtac	(ON_SHC349)	(SEQ ID NO:28)
40	4	VHSzy9	ATGtattactgtgc	(ON_SHC5a)	(SEQ ID NO:33)

Table 12D, testing 79 human HC V genes with four probes

Number of sequences..... 79  
Number of bases..... 29143

		Number of mismatches								
	Id	Best	0	1	2	3	4	5		
5	1	39	15	11	10	1	2	0	Seq1	gtgtattactgtgc (SEQ ID NO:25)
	2	22	7	6	5	3	0	1	Seq2	gtAtattactgtgc (SEQ ID NO:26)
	3	7	1	5	1	0	0	0	Seq4	gtgtattactgtAc (SEQ ID NO:28)
	4	11	2	4	4	1	0	0	Seq9	ATgtattactgtgc (SEQ ID NO:33)
	Group		25	26	20	5	2			
10	Cumulative		25	51	71	76	78			

One sequence has five mismatches with sequences 2, 4, and 9; it is scored as best for 2.

- Id is the number of the adapter.  
 Best is the number of sequence for which the identified adapter was the best available.
- 15 The rest of the table shows how well the sequences match the adapters. For example, there are 10 sequences that match VHSzyl(Id=1) with 2 mismatches and are worse for all other adapters. In this sample, 90% come within 2 bases of one of
- 20 the four adapters.

Table 13

The following list of enzymes was taken from  
<http://rebase.neb.com/cgi-bin/asymmlist>.

- 5 I have removed the enzymes that a) cut within the recognition, b) cut on both sides of the recognition, or c) have fewer than 2 bases between recognition and closest cut site.

REBASE Enzymes  
 04/13/2001

10	Type II restriction enzymes with asymmetric recognition sequences:			
	Enzymes	Recognition Sequence	Isoschizomers	Suppliers
	AarI	CACCTGCNNNN^NNNN	-	Y
	AceIII	CAGCTCNNNNNNN^NNNN	-	-
	Bbr7I	GAAGACNNNNNNN^NNNN	-	-
15	BbvI	GCAGCNNNNNNN^NNNN	-	Y
	BbvII	GAAGACNN^NNNN	-	-
	Bce83I	CTTGAGNNNNNNNNNNNNNN_NN^	-	-
	BceAI	ACGGCNNNNNNNNNNNN^NN	-	Y
	BceFI	ACGGCNNNNNNNNNNNN^N	-	-
20	BciVI	GTATCCNNNNN N^	BfuI	Y
	BfiI	ACTGGGNNNN N^	BmrI	Y
	BinI	GGATCNNNN^N	-	-
	BscAI	GCATCNNNN^NN	-	-
	BseRI	GAGGAGNNNNNNNN_NN^	-	Y
25	BsmFI	GGGACNNNNNNNNNN^NNNN	BspLU11III	Y
	BspMI	ACCTGCNNNN^NNNN	Acc36I	Y
	EciI	GGCGGANNNNNNNNNN_NN^	-	Y
	Eco57I	CTGAAGNNNNNNNNNNNNNN_NN^	BspKT5I	Y
	FauI	CCCGCNNNN^NN	BstFZ438I	Y
30	FokI	GGATGNNNNNNNNN^NNNN	BstPZ418I	Y
	GsuI	CTGGAGNNNNNNNNNNNNNN_NN^	-	Y
	HgaI	GACGCNNNNN^NNNNN	-	Y
	HphI	GGTGANNNNNNN N^	AsuHPI	Y
	MboII	GAAGANNNNNNN N^	-	Y
35	MlyI	GAGTCNNNNN^	SchI	Y
	MmeI	TCCRACNNNNNNNNNNNNNNNN_NN^	-	-
	MnlI	CCTCNNNNNN N^	-	Y
	PleI	GAGTCNNNN^N	PpsI	Y
	RleAI	CCCACANNNNNNNNN_NNN^	-	-
40	SfaNI	GCATCNNNNN^NNNN	BspST5I	Y
	SspD5I	GGTGANNNNNNNN^	-	-
	Sth132I	CCCGNNNN^NNNN	-	-
	StsI	GGATGNNNNNNNNNN^NNNN	-	-
	TaqII	GACCGANNNNNNNNN_NN^, CACCCANNNNNNNNN_NN^	-	-
45	Tth111II	CAARCAANNNNNNNNN_NN^	-	-
	UbaPI	CGAACG	-	-

The notation is ^ means cut the upper strand and \_ means cut the lower strand. If the upper and lower strand are cut at the same place, then only ^ appears.





Table 15: Use of *FokI* as "Universal Restriction Enzyme"

*FokI* - for dsDNA, | represents sites of cleavage

5' - cacGGATGtg--nnnnnnnn|nnnnnnnn-3' (SEQ ID NO:15)  
 3' - gtgCCTACac--nnnnnnnnnn|nnn-5' (SEQ ID NO:16)  
 RECOG  
 NITion of *FokI*

Case I

5' - ...gtg|tatt-actgtgc..Substrate....-3' (SEQ ID NO:17)  
 3' - cac-ataa|tgacacg -  
 gtGTAGGcac\  
 5' - caCATCCgtg/ (SEQ ID NO:18)

Case II

5' - ...gtgtatt|agac-tgc..Substrate....-3' (SEQ ID NO:19)  
 -cacataa-tctg|acg-5'  
 /gtgCCTACac  
 \cacGGATGtg-3' (SEQ ID NO:20)

Case III (Case I rotated 180 degrees)

/gtgCCTACac-5'  
 \cacGGATGtg-  
 gtgtctt|acag-tcc-3' Adapter (SEQ ID NO:21)  
 3' - ...cacagaa-tgtc|agg..substrate....-5' (SEQ ID NO:22)

Case IV (Case II rotated 180 degrees)

3' - gtGTAGGcac\ (SEQ ID NO:23)  
 -caCATCCgtg/  
 5' - gag|tctc-actgagc  
 Substrate 3' - ...ctc-agag|tgactcg...-5' (SEQ ID NO:24)

Improved *FokI* adapters

*FokI* - for dsDNA, | represents sites of cleavage

Case I

Stem 11, loop 5, stem 11, recognition 17

5' - ...catgtg|tatt-actgtgc..Substrate....-3'  
 3' - gtacac-ataa|tgacacg -  
 gtGTAGGcacG T  
 5' - caCATCCgtgc C  
 [TT]

Case II

Stem 10, loop 5, stem 10, recognition 18

5'...gtgtatt|agac-tgctgcc..Substrate....-3'  
 5     ┌T┐     ┌cacataa-tctg|acgacgg-5'  
      T   gtgCCTACac  
      C   cacGGATGtg-3'  
      └T┘

Case III (Case I rotated 180 degrees)

Stem 11, loop 5, stem 11, recognition 20

10     ┌T┐  
      T   TgtgCCTACac-5'  
      G   AcacGGATGtg┐  
      └T┘     gtgtctt|acag-tccattctg-3' Adapter  
              3'...cacagaa-tgtc|aggtaagac..substrate....-5'

15 Case IV (Case II rotated 180 degrees)

Stem 11, loop 4, stem 11, recognition 17

  ┌T┐  
   3'- gtGTAGGcacc T  
   ┌caCATCCgtgg T  
 20     5'-atcgag|tctc-actgagc     └T┘  
      Substrate 3'...tagctc-agag|tgactcg...-5'

**BseRI**

  | sites of cleavage  
 25     5'-cacGAGGAGnnnnnnnnnn|nnnnn-3'  
       3'-gtgctcctcnnnnnnnnn|nnnnnn-5'  
           RECOG  
           NITion of BseRI

Stem 11, loop 5, stem 11, recognition 19

3'-.....gaacat|cg-ttaagccagta.....5'  
 30     ┌T-T┐     cttgta-gc|aattcggtcat-3'  
      C   GCTGAGGAGTC-┐  
      T   cgactcctcag-5' An adapter for BseRI to cleave the substrate above.  
      └T┘

Table 16 Human heavy chains bases 88.1 to 94.2

Number of sequences..... 840

5	Id	Ntot	Number of Mismatches.....							Name	Probe	Sequence.....	Dot form.....
			0	1	2	3	4	5	6	7			
10	1	364	152	97	76	26	7	4	2	0	VHS881-1.1	gctgtgtattactgtgcgag	gctgtgtattactgtgcgag
	2	265	150	60	33	13	5	4	0	0	VHS881-1.2	gccgtgtattactgtgcgag	..C.....
	3	96	14	34	16	10	5	7	9	1	VHS881-2.1	gccgtatattactgtgcgag	..C.a.....
	4	20	0	3	4	9	2	2	0	0	VHS881-4.1	gccgtgtattactgtacgag	..C.....a...
	5	95	25	36	18	11	2	2	0	1	VHS881-9.1	gccatgtattactgtgcgag	..ca.....

840 341 230 147 69 21 19 11 2  
341 571 718 787 808 827 838 840

88 89 90 91 92 93 94 95 Codon number as in Table 195

Recognition..... Stem..... Loop. Stem.....  
(VHS881-1.1) 5'-gctgtgtat|tact-gtgcgag cAcATccgTg TTgTT cAcgATgTg-3'  
(VHS881-1.2) 5'-gccgtgtat|tact-gtgcgag cAcATccgTg TTgTT cAcgATgTg-3'  
(VHS881-2.1) 5'-gccgtgtat|tact-gtgcgag cAcATccgTg TTgTT cAcgATgTg-3'  
(VHS881-4.1) 5'-gccgtgtat|tact-gtgcgag cAcATccgTg TTgTT cAcgATgTg-3'  
(VHS881-9.1) 5'-gccatgtat|tact-gtgcgag cAcATccgTg TTgTT cAcgATgTg-3'

| site of substrate cleavage

(FOKIact) 5'-cAcATccgTg TTgTT cAcgATgTg-3'

(VHEX881) 5'-AATAgTAgAc TgcAgTgTcc TcAgcccTTA AgcTgTTcAT cTgcAAGTAg-  
AgAgTATTCT TAGAgTTgTc TcTAGAcTTA gTgAAGg-3'

! note that VHEX881 is the reverse complement of the ON below

25 ! [RC] 5'-cgCttcacTaa-

! Scab.....

! Synthetic 3-23 as in Table 206

! |TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-

! XbaI...

- 186 -

! |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|t-3'  
! AHII...  
(VHBA881) 5'-cgCttcacTaa-  
|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-  
|aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgt gcg ag-3'  
(VHBB881) 5'-cgCttcacTaa-  
|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-  
|aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgt Acg ag-3'  
(VH881PCR) 5'-cgCttcacTaa|TCT|AGA|gac|aac-3'



FokI. FokI.

What happens in the upper strand:

(SzKB1230-O12\*) 5'-gac cca gtc | tcc a-tc ctc c-3'  
| Site of cleavage in substrate

5 (SzKB1230-A17\*) 5'-gac tca gtc | tcc a-ct ctc c-3'

(SzKB1230-A27\*) 5'-gac gca gtc | tcc a-gg cac c-3'

10 (SzKB1230-A11\*) 5'-gac gca gtc | tcc a-gc cac c-3'

(kapextURE) 5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg-3' !sense strand  
Scab.....ApaLI.

(kapextUREPCR) 5' -ccTctactctTgTcAcAgTg-3'  
Scab.....

15 (kaBRO1UR) 5' -ggAggATggA cTggATgTcT' TgTgcAcTgT' gAcAAGAgTA gAgg-3'  
[RC] 5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg tcc a-tc ctc c-3' ON above is R.C. of this one  
(kaBRO2UR) 5' -ggAggATggA cTggATgTcT' TgTgcAcTgT' gAcAAGAgTA gAgg-3'  
[RC] 5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg tcc a-ct ctc c-3' ON above is R.C. of this one  
(kaBRO3UR) 5' -ggTgcTggA cTggATgTcT' TgTgcAcTgT' gAcAAGAgTA gAgg-3'  
[RC] 5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg tcc a-gg cac c-3' ON above is R.C. of this one  
(kaBRO4UR) 5' -ggTggcTggA cTggATgTcT' TgTgcAcTgT' gAcAAGAgTA gAgg-3'  
[RC] 5' -ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg tcc a-gc cac c-3' ON above is R.C. of this one  
Scab.....ApaLI.









Table 19: Cleavage of 75 human light chains.

	Enzyme	Recognition*	Nch	Ns	Planned location of site
	AfeI	AGCgct	0	0	
5	<b>AflIII</b>	<b>Cttaag</b>	<b>0</b>	<b>0</b>	<b>HC FR3</b>
	AgeI	Accggt	0	0	
	<b>AscI</b>	<b>GGcgcgcc</b>	<b>0</b>	<b>0</b>	<b>After LC</b>
	BglII	Agatct	0	0	
	BsiWI	Cgtacg	0	0	
	BspDI	ATcgat	0	0	
10	BssHII	Gcgcg	0	0	
	BstBI	TTcgaa	0	0	
	DraIII	CACNNNgtg	0	0	
	EagI	Cggccg	0	0	
	FseI	GGCCGGcc	0	0	
15	FspI	TGCgca	0	0	
	HpaI	GTTaac	0	0	
	<b>MfeI</b>	<b>Caattg</b>	<b>0</b>	<b>0</b>	<b>HC FR1</b>
	MluI	Acgcgt	0	0	
	<b>NcoI</b>	<b>Ccatgg</b>	<b>0</b>	<b>0</b>	<b>Heavy chain signal</b>
20	<b>NheI</b>	<b>Gctagc</b>	<b>0</b>	<b>0</b>	<b>HC/anchor linker</b>
	<b>NotI</b>	<b>GCggccgc</b>	<b>0</b>	<b>0</b>	<b>In linker after HC</b>
	NruI	TCGcga	0	0	
	PacI	TTAATtaa	0	0	
	PmeI	GTTTaaac	0	0	
25	PmlI	CACgtg	0	0	
	PvuI	CGATcg	0	0	
	SacII	CCGCgg	0	0	
	SalI	Gtcgac	0	0	
	<b>SfiI</b>	<b>GGCCNNNNnggcc</b>	<b>0</b>	<b>0</b>	<b>Heavy Chain signal</b>
30	SgfI	GCGATcg	0	0	
	SnaBI	TACgta	0	0	
	StuI	AGGcct	0	0	
	<b>XbaI</b>	<b>Tctaga</b>	<b>0</b>	<b>0</b>	<b>HC FR3</b>
	AatII	GACGTc	1	1	
35	AclI	AACgtt	1	1	
	AseI	ATtaat	1	1	
	BsmI	GAATGCN	1	1	
	<b>BspEI</b>	<b>Tccgga</b>	<b>1</b>	<b>1</b>	<b>HC FR1</b>
	<b>BstXI</b>	<b>CCANNNNNntgg</b>	<b>1</b>	<b>1</b>	<b>HC FR2</b>
40	DrdI	GACNNNNnngtc	1	1	
	HindIII	Aagctt	1	1	
	PciI	Acatgt	1	1	
	SapI	gaagagc	1	1	
	ScaI	AGTact	1	1	
45	SexAI	Accwgg	1	1	
	SpeI	Actagt	1	1	
	TliI	Ctcgag	1	1	
	XhoI	Ctcgag	1	1	
	BcgI	cgannnnnntgc	2	2	
50	BlpI	GCTnagc	2	2	
	BssSI	Ctcgtg	2	2	
	BstAPI	GCANNNNntgc	2	2	
	EspI	GCTnagc	2	2	
	KasI	Ggcgcc	2	2	
55	PflMI	CCANNNNntgg	2	2	
	XmnI	GAANNnttc	2	2	
	<b>ApaLI</b>	<b>Gtgcac</b>	<b>3</b>	<b>3</b>	<b>LC signal seq</b>

	NaeI	GCCggc	3	3	
	NgoMI	Gccggc	3	3	
	PvuII	CAGctg	3	3	
	RsrII	CGGwccg	3	3	
5	BsrBI	GAGcgg	4	4	
	BsrDI	GCAATGNNn	4	4	
	BstZ17I	GTAtac	4	4	
	EcoRI	Gaattc	4	4	
	SphI	GCATGc	4	4	
10	SspI	AATatt	4	4	
	AccI	GTmkac	5	5	
	BclI	Tgatca	5	5	
	BsmBI	Nnnnnngagacg	5	5	
	BsrGI	Tgtaca	5	5	
15	DraI	TTTaaa	6	6	
	<b>NdeI</b>	<b>CAtatg</b>	<b>6</b>	<b>6</b>	<b>HC FR4</b>
	SwaI	ATTTaaat	6	6	
	BamHI	Ggatcc	7	7	
	SacI	GAGCTc	7	7	
20	BciVI	GTATCCNNNNNN	8	8	
	BsaBI	GATNNnnatc	8	8	
	NsiI	ATGCAc	8	8	
	<b>Bsp120I</b>	<b>Gggccc</b>	<b>9</b>	<b>9</b>	<b>CH1</b>
	<b>ApaI</b>	<b>GGGCCc</b>	<b>9</b>	<b>9</b>	<b>CH1</b>
25	PspOoMI	Gggccc	9	9	
	BspHI	Tcatga	9	11	
	EcoRV	GATatc	9	9	
	AhdI	GACNNNngtc	11	11	
	BbsI	GAAGAC	11	14	
30	PsiI	TTAtaa	12	12	
	BsaI	GGTCTCnNNnn	13	15	
	XmaI	Cccggg	13	14	
	AvaI	Cycgrg	14	16	
	BglI	GCCNNNNnggc	14	17	
35	AlwNI	CAGNNNctg	16	16	
	BspMI	ACCTGC	17	19	
	XcmI	CCANNNNNnnnnntgg	17	26	
	<b>BstEII</b>	<b>Ggtnacc</b>	<b>19</b>	<b>22</b>	<b>HC FR4</b>
	Sse8387I	CCTGCAGg	20	20	
40	AvrII	Cctagg	22	22	
	HincII	GTYrac	22	22	
	BsgI	GTGCAG	27	29	
	MscI	TGGcca	30	34	
	BseRI	NNnnnnnnnnctcctc	32	35	
45	Bsu36I	CCtnagg	35	37	
	PstI	CTGCAG	35	40	
	EciI	nnnnnnnnntccgcc	38	40	
	PpuMI	RGgwccy	41	50	
	StyI	Ccwwgg	44	73	
50	EcoO109I	RGgnccy	46	70	
	Acc65I	Ggtacc	50	51	
	KpnI	GGTACc	50	51	
	BpmI	ctccag	53	82	
	AvaII	Ggwcc	71	124	

55 \* cleavage occurs in the top strand after the last upper-case base. For REs that cut palindromic sequences, the lower strand is cut at the symmetrical site.

Table 20: Cleavage of 79 human heavy chains

	Enzyme	Recognition	Nch	Ns	Planned location of site
	AfeI	AGCgct	0	0	
5	<b>AflIII</b>	<b>Cttaag</b>	<b>0</b>	<b>0</b>	<b>HC FR3</b>
	<b>AscI</b>	<b>GGGgagcc</b>	<b>0</b>	<b>0</b>	<b>After LC</b>
	BsiWI	Cgtacg	0	0	
	BspDI	ATcgat	0	0	
	BssHII	Gcgcg	0	0	
10	FseI	GGCCGGcc	0	0	
	HpaI	GTTaac	0	0	
	<b>NheI</b>	<b>Gctagc</b>	<b>0</b>	<b>0</b>	<b>HC Linker</b>
	<b>NotI</b>	<b>GCggccgc</b>	<b>0</b>	<b>0</b>	<b>In linker, HC/anchor</b>
	NruI	TCGcga	0	0	
	NsiI	ATGCAT	0	0	
15	PacI	TTAATtaa	0	0	
	PciI	Acatgt	0	0	
	PmeI	GTTTaaac	0	0	
	PvuI	CGATcg	0	0	
20	RsrII	CGgwccg	0	0	
	SapI	gaagagc	0	0	
	<b>SfiI</b>	<b>GGCCNNNNnggcc</b>	<b>0</b>	<b>0</b>	<b>HC signal seq</b>
	SgfI	GCGATcgc	0	0	
	SwaI	ATTTaaat	0	0	
25	AclI	AACgtt	1	1	
	AgeI	Accggt	1	1	
	AseI	ATtaat	1	1	
	AvrII	Cctagg	1	1	
	BsmI	GAATGCN	1	1	
	BsrBI	GAGcgg	1	1	
30	BsrDI	GCAATGNNn	1	1	
	DraI	TTTaaa	1	1	
	EspI	TGCgca	1	1	
	HindIII	Aagctt	1	1	
35	<b>MfeI</b>	<b>Caattg</b>	<b>1</b>	<b>1</b>	<b>HC FR1</b>
	NaeI	GCCggc	1	1	
	NgoMI	Gccggc	1	1	
	SpeI	Actagt	1	1	
	Acc65I	Ggtacc	2	2	
40	BstBI	TTcgaa	2	2	
	KpnI	GGTACc	2	2	
	MluI	Acgcgt	2	2	
	<b>NcoI</b>	<b>Ccatgg</b>	<b>2</b>	<b>2</b>	<b>In HC signal seq</b>
	<b>NdeI</b>	<b>CAtatg</b>	<b>2</b>	<b>2</b>	<b>HC FR4</b>
	PmlI	CACgtg	2	2	
45	XcmI	CCANNNNNnnntgg	2	2	
	BcgI	cgannnnntgc	3	3	
	BclI	Tgatca	3	3	
	BglI	GCCNNNNnggc	3	3	
	BsaBI	GATNNnnatc	3	3	
50	BsrGI	Tgtaca	3	3	
	SnaBI	TACgta	3	3	
	Sse8387I	CCTGCagg	3	3	
	<b>ApaLI</b>	<b>Gtgcac</b>	<b>4</b>	<b>4</b>	<b>LC Signal/FR1</b>
55	BspHI	Tcatga	4	4	
	BssSI	Ctcgtg	4	4	
	PsiI	TTAtaa	4	5	

Table 20: Cleavage of 79 human heavy chains

	SphI	GCATGc	4	4	
	AhdI	GACNNNnngtc	5	5	
	<b>BspEI</b>	<b>Tccgga</b>	<b>5</b>	<b>5</b>	<b>HC FR1</b>
5	MscI	TGGcca	5	5	
	SacI	GAGCTc	5	5	
	ScaI	AGTact	5	5	
	SexAI	Accwgg	5	6	
	SspI	AATatt	5	5	
	TliI	Ctcgag	5	5	
10	XhoI	Ctcgag	5	5	
	BbsI	GAAGAC	7	8	
	BstAPI	GCANNNNntgc	7	8	
	BstZ17I	GTAtac	7	7	
	EcoRV	GATatc	7	7	
15	EcoRI	Gaattc	8	8	
	BlpI	GCtnagc	9	9	
	Bsu36I	CCtnagg	9	9	
	DraIII	CACNNNgtg	9	9	
	EspI	GCtnagc	9	9	
20	StuI	AGGcct	9	13	
	<b>XbaI</b>	<b>Tctaga</b>	<b>9</b>	<b>9</b>	<b>HC FR3</b>
	<b>Bsp120I</b>	<b>Gggccc</b>	<b>10</b>	<b>11</b>	<b>CH1</b>
	<b>ApaI</b>	<b>GGGCCc</b>	<b>10</b>	<b>11</b>	<b>CH1</b>
	PspOoMI	Gggccc	10	11	
25	BciVI	GTATCCNNNNNN	11	11	
	SalI	Gtcgac	11	12	
	DrdI	GACNNNNnngtc	12	12	
	KasI	Ggcgcc	12	12	
	XmaI	Cccggg	12	14	
30	BglII	Agatct	14	14	
	HincII	GTYrac	16	18	
	BamHI	Ggatcc	17	17	
	PflMI	CCANNNNntgg	17	18	
	BsmBI	Nnnnnngagacg	18	21	
35	<b>BstXI</b>	<b>CCANNNNntgg</b>	<b>18</b>	<b>19</b>	<b>HC FR2</b>
	XmnI	GAANNnnttc	18	18	
	SacII	CCGCgg	19	19	
	PstI	CTGCAG	20	24	
	PvuII	CAGctg	20	22	
40	AvaI	Cycgrg	21	24	
	EagI	Cggccg	21	22	
	AatII	GACGTc	22	22	
	BspMI	ACCTGC	27	33	
	AccI	GTmkac	30	43	
45	StyI	Ccwwgg	36	49	
	AlwNI	CAGNNNctg	38	44	
	BsaI	GGTCTCNnnnn	38	44	
	PpuMI	RGgwccy	43	46	
	BsgI	GTGCAG	44	54	
50	BseRI	NNnnnnnnnnctctc	48	60	
	EciI	nnnnnnnnntccgcc	52	57	
	<b>BstEII</b>	<b>Ggtnacc</b>	<b>54</b>	<b>61</b>	<b>HC Fr4, 47/79 have one</b>
	EcoO109I	RGgnccy	54	86	
	BpmI	ctccag	60	121	
55	AvaII	Ggwcc	71	140	

Table 21: MALIA3, annotated

! MALIA3 9532 bases

```

-----
5  !   1 aat gct act act att agt aga att gat gcc acc ttt tca gct cgc gcc
      gene ii continued
      49 cca aat gaa aat ata gct aaa cag gtt att gac cat ttg cga aat gta
      97 tct aat ggt caa act aaa tct act cgt tcg cag aat tgg gaa tca act
      145 gtt aca tgg aat gaa act tcc aga cac cgt act tta gtt gca tat tta
10  !   193 aaa cat gtt gag cta cag cac cag att cag caa tta agc tct aag cca
      241 tcc gca aaa atg acc tct tat caa aag gag caa tta aag gta ctc tct
      289 aat cct gac ctg ttg gag ttt gct tcc ggt ctg gtt cgc ttt gaa gct
      337 cga att aaa acg cga tat ttg aag tct ttc ggg ctt cct ctt aat ctt
      385 ttt gat gca atc cgc ttt gct tct gac tat aat agt cag ggt aaa gac
      433 ctg att ttt gat tta tgg tca ttc tcg ttt tct gaa ctg ttt aaa gca
15  !   481 ttt gag ggg gat tca ATG aat att tat gac gat tcc gca gta ttg gac
      RBS?..... Start gene x, ii continues
      529 gct atc cag tct aaa cat ttt act att acc ccc tct ggc aaa act tct
      577 ttt gca aaa gcc tct cgc tat ttt ggt ttt tat cgt cgt ctg gta aac
      625 gag ggt tat gat agt gtt gct ctt act atg cct cgt aat tcc ttt tgg
20  !   673 cgt tat gta tct gca tta gtt gaa tgt ggt att cct aaa tct caa ctg
      721 atg aat ctt tct acc tgt aat aat gtt gtt ccg tta gtt cgt ttt att
      769 aac gta gat ttt tct tcc caa cgt cct gac tgg tat aat gag cca gtt
      817 ctt aaa atc gca TAA
      End X & II
25  !   832 ggtaattca ca
      !
      !           M1           E5           Q10           T15
      843 ATG att aaa gtt gaa att aaa cca tct caa gcc caa ttt act act cgt
      Start gene V
30  !
      !           S17           S20           P25           E30
      891 tct ggt gtt tct cgt cag ggc aag cct tat tca ctg aat gag cag ctt
      !
      !           V35           E40           V45
      939 tgt tac gtt gat ttg ggt aat gaa tat ccg gtt ctt gtc aag att act
35  !
      !           D50           A55           L60
      987 ctt gat gaa ggt cag cca gcc tat gcg cct ggt cTG TAC Acc gtt cat
      BsrGI...
40  !           L65           V70           S75           R80
      1035 ctg tcc tct ttc aaa gtt ggt cag ttc ggt tcc ctt atg att gac cgt
      !
      !           P85           K87 end of V
      1083 ctg cgc ctc gtt ccg gct aag TAA C
45  !
      !   1108 ATG gag cag gtc gcg gat ttc gac aca att tat cag gcg atg
      Start gene VII
      !
      !   1150 ata caa atc tcc gtt gta ctt tgt ttc gcg ctt ggt ata atc
50  !
      !           VII and IX overlap.
      !           ..... S2 V3 L4 V5           S10
      1192 gct ggg ggt caa agA TGA gt gtt tta gtg tat tct ttc gcc tct ttc gtt
      !
      !           End VII
      !           |start IX
55  !           L13           W15           G20           T25           E29
      1242 tta ggt tgg tgc ctt cgt agt ggc att acg tat ttt acc cgt tta atg gaa

```

```

!
1293 act tcc tc
!
!      .... stop of IX, IX and VIII overlap by four bases
5 1301 ATG aaa aag tct tta gtc ctc aaa gcc tct gta gcc gtt gct acc ctc
!      Start signal sequence of viii.
!
1349 gtt ccg atg ctg tct ttc gct gct gag ggt gac gat ccc gca aaa gcg
!      mature VIII --->
10 1397 gcc ttt aac tcc ctg caa gcc tca gcg acc gaa tat atc ggt tat gcg
1445 tgg gcg atg gtt gtt gtc att
1466 gtc ggc gca act atc ggt atc aag ctg ttt aag
1499 aaa ttc acc tcg aaa gca ! 1515
!      ..... -35 ..
15
1517 agc tga taaaccgat acaattaaag gctccttttg
!      ..... -10 ...
!
1552 gagccttttt ttttGGAGAt ttt ! S.D. underlined
20
!      <----- III signal sequence ----->
!      M K K L L F A I P L V
1575 caac GTG aaa aaa tta tta ttc gca att cct tta gtt ! 1611
!
25 V P F Y S H S A Q
1612 gtt cct ttc tat tct cac aGT gcA Cag tCT
!      ApaLI...
!
1642 GTC GTG ACG CAG CCG CCC TCA GTG TCT GGG GCC CCA GGG CAG
30 AGG GTC ACC ATC TCC TGC ACT GGG AGC AGC TCC AAC ATC GGG GCA
!      BstEII...
1729 GGT TAT GAT GTA CAC TGG TAC CAG CAG CTT CCA GGA ACA GCC CCC AAA
1777 CTC CTC ATC TAT GGT AAC AGC AAT CGG CCC TCA GGG GTC CCT GAC CGA
1825 TTC TCT GGC TCC AAG TCT GGC ACC TCA GCC TCC CTG GCC ATC ACT
35 1870 GGG CTC CAG GCT GAG GAT GAG GCT GAT TAT
1900 TAC TGC CAG TCC TAT GAC AGC AGC CTG AGT
1930 GGC CTT TAT GTC TTC GGA ACT GGG ACC AAG GTC ACC GTC
!      BstEII...
1969 CTA GGT CAG CCC AAG GCC AAC CCC ACT GTC ACT
40 2002 CTG TTC CCG CCC TCC TCT GAG GAG CTC CAA GCC AAC AAG GCC ACA CTA
2050 GTG TGT CTG ATC AGT GAC TTC TAC CCG GGA GCT GTG ACA GTG GCC TGG
2098 AAG GCA GAT AGC AGC CCC GTC AAG GCG GGA GTG GAG ACC ACC ACA CCC
2146 TCC AAA CAA AGC AAC AAC AAG TAC GCG GCC AGC AGC TAT CTG AGC CTG
2194 ACG CCT GAG CAG TGG AAG TCC CAC AGA AGC TAC AGC TGC CAG GTC ACG
45 2242 CAT GAA GGG AGC ACC GTG GAG AAG ACA GTG GCC CCT ACA GAA TGT TCA
2290 TAA TAA ACCG CCTCCACCGG GCGCGCCAAT TCTATTTCAA GGAGACAGTC ATA
!      AscI.....
!
!      PelB signal----->
50 M K Y L L P T A A A G L L L L
1343 ATG AAA TAC CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC
!
!      16 17 18 19 20 21 22
!      A A Q P A M A
55 2388 gcG GCC cag ccG GCC atg gcc
!      SfiI.....
!      NgoMI...(1/2)
!      NcoI.....

```

1004074 "102501"





```

!           A   S   T   K   G   P   S   V   F   P
2769      gcc tcc acc aaG GGC CCa tcg GTC TTC ccc
!           Bsp120I.      BbsI...(2/2)
!           ApaI....
5
!
!       153 154 155 156 157 158 159 160 161 162 163 164 165 166 167
!       L   A   P   S   S   K   S   T   S   G   G   T   A   A   L
2799      ctg gca ccC TCC TCc aag agc acc tct ggg ggc aca gcg gcc ctg
!       BseRI...(2/2)
10
!
!       168 169 170 171 172 173 174 175 176 177 178 179 180 181 182
!       G   C   L   V   K   D   Y   F   P   E   P   V   T   V   S
2844      ggc tgc ctg GTC AAG GAC TAC TTC CCc gaA CCG GTg acg gtg tcg
!       AgeI....
15
!
!       183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
!       W   N   S   G   A   L   T   S   G   V   H   T   F   P   A
2889      tgg aac tca GGC GCC ctg acc agc ggc gtc cac acc ttc ccg gct
!       KasI...(1/4)
20
!
!       198 199 200 201 202 203 204 205 206 207 208 209 210 211 212
!       V   L   Q   S   S   G   L   Y   S   L   S   S   V   V   T
2934      gtc cta cag tCt agc GGa ctc tac tcc ctc agc agc gta gtg acc
!       (Bsu36I...)(knocked out)
25
!
!       213 214 215 216 217 218 219 220 221 222 223 224 225 226 227
!       V   P   S   S   S   L   G   T   Q   T   Y   I   C   N   V
2979      gtg ccC tCt tct agc tTG Ggc acc cag acc tac atc tgc aac gtg
!       (BstXI.....)N.B. destruction of BstXI & BpmI sites.
30
!
!       228 229 230 231 232 233 234 235 236 237 238 239 240 241 242
!       N   H   K   P   S   N   T   K   V   D   K   K   V   E   P
3024      aat cac aag ccc agc aac acc aag gtg gac aag aaa gtt gag ccc
!
!       243 244 245
!       K   S   C   A   A   A   H   H   H   H   H   H   S   A
3069      aaa tct tgt GCG GCC GcT cat cac cac cat cat cac tct gct
!       NotI.....
35
!
!       40
!       3111      gaa caa aaa ctc atc tca gaa gag gat ctg aat ggt gcc gca
!
!       D   I   N   D   D   R   M   A   S   G   A
45      3153      GAT ATC aac gat gat cgt atg   gct AGC   ggc gcc
!       rEK cleavage site..... NheI... KasI...
!       EcoRV..
!
!       Domain 1 -----
50      3183      gct gaa act gtt gaa agt tgt tta gca
!
!       K   P   H   T   E   I   S   F
55      3210      aaa ccc cat aca gaa aat tca ttt
!
!       T   N   V   W   K   D   D   K   T
3234      acT AAC GTC TGG AAA GAC GAC AAA ACT

```

10045674-100504



```

!       S   V   E   C   R   P   F   V   F   S   A   G   K   P   Y   E
4233 tcg gtt gaa tgt cgc cct ttt gtc ttt agc gct ggt aaa cca tat gaa
!
!       F   S   I   D   C   D   K   I   N   L   F   R
5 4281 ttt tct att gat tgt gac aaa ata aac tta ttc cgt
!                                     End Domain 3
!
!       G   V   F   A   F   L   L   Y   V   A   T   F   M   Y   V   F140
10 4317 ggt gtc ttt gcg ttt ctt tta tat gtt gcc acc ttt atg tat gta ttt
!       start transmembrane segment
!
!       S   T   F   A   N   I   L
4365 tct acg ttt gct aac ata ctg
!
!       R   N   K   E   S
15 4386 cgt aat aag gag tct TAA ! stop of iii
!       Intracellular anchor.
!
!       M1 P2 V L L5 G I P L L10 L R F L G15
20 4404 tc ATG cca gtt ctt ttg ggt att ccg tta tta ttg cgt ttc ctc ggt
!       Start VI
!
4451 ttc ctt ctg gta act ttg ttc ggc tat ctg ctt act ttt ctt aaa aag
4499 ggc ttc ggt aag ata gct att gct att tca ttg ttt ctt gct ctt att
25 4547 att ggg ctt aac tca att ctt gtg ggt tat ctc tct gat att agc gct
4595 caa tta ccc tct gac ttt gtt cag ggt gtt cag tta att ctc ccg tct
4643 aat gcg ctt ccc tgt ttt tat gtt att ctc tct gta aag gct gct att
4691 ttc att ttt gac gtt aaa caa aaa atc gtt tct tat ttg gat tgg gat
!
!       M1 A2 V3 F5 L10 G13
30 4739 aaa TAA t ATG gct gtt tat ttt gta act ggc aaa tta ggc tct gga
!       end VI Start gene I
!
!       14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
35 4785 aag acg ctc gtt agc gtt ggt aag att cag gat aaa att gta gct
!
!       29 30 31 32 33 34 35 36 37 38 39 40 41 42 43
40 4830 ggg tgc aaa ata gca act aat ctt gat tta agg ctt caa aac ctc
!
!       44 45 46 47 48 49 50 51 52 53 54 55 56 57 58
45 4875 ccg caa gtc ggg agg ttc gct aaa acg cct cgc gtt ctt aga ata
!
!       59 60 61 62 63 64 65 66 67 68 69 70 71 72 73
4920 ccg gat aag cct tct ata tct gat ttg ctt gct att ggg cgc ggt
!
!       74 75 76 77 78 79 80 81 82 83 84 85 86 87 88
50 4965 aat gat tcc tac gat gaa aat aaa aac ggc ttg ctt gtt ctc gat
!
!       89 90 91 92 93 94 95 96 97 98 99 100 101 102 103
55 5010 gag tgc ggt act tgg ttt aat acc cgt tct tgg aat gat aag gaa
!
!       104 105 106 107 108 109 110 111 112 113 114 115 116 117 118

```

10045674-1005004

```

!      R   Q   P   I   I   D   W   F   L   H   A   R   K   L   G
5055  aga  cag  ccg  att  att  gat  tgg  ttt  cta  cat  gct  cgt  aaa  tta  gga
!
!      119 120 121 122 123 124 125 126 127 128 129 130 131 132 133
5      W   D   I   I   F   L   V   Q   D   L   S   I   V   D   K
!
5100  tgg  gat  att  att  ttt  ctt  gtt  cag  gac  tta  tct  att  gtt  gat  aaa
!
!      134 135 136 137 138 139 140 141 142 143 144 145 146 147 148
10     Q   A   R   S   A   L   A   E   H   V   V   Y   C   R   R
!
5145  cag  gcg  cgt  tct  gca  tta  gct  gaa  cat  gtt  gtt  tat  tgt  cgt  cgt
!
!      149 150 151 152 153 154 155 156 157 158 159 160 161 162 163
15     L   D   R   I   T   L   P   F   V   G   T   L   Y   S   L
!
5190  ctg  gac  aga  att  act  tta  cct  ttt  gtc  ggt  act  tta  tat  tct  ctt
!
!      164 165 166 167 168 169 170 171 172 173 174 175 176 177 178
!      I   T   G   S   K   M   P   L   P   K   L   H   V   G   V
!
5235  att  act  ggc  tcg  aaa  atg  cct  ctg  cct  aaa  tta  cat  gtt  ggc  gtt
!
!      179 180 181 182 183 184 185 186 187 188 189 190 191 192 193
20     V   K   Y   G   D   S   Q   L   S   P   T   V   E   R   W
!
5280  gtt  aaa  tat  ggc  gat  tct  caa  tta  agc  cct  act  gtt  gag  cgt  tgg
!
!      194 195 196 197 198 199 200 201 202 203 204 205 206 207 208
25     L   Y   T   G   K   N   L   Y   N   A   Y   D   T   K   Q
!
5325  ctt  tat  act  ggt  aag  aat  ttg  tat  aac  gca  tat  gat  act  aaa  cag
!
!      209 210 211 212 213 214 215 216 217 218 219 220 221 222 223
30     A   F   S   S   N   Y   D   S   G   V   Y   S   Y   L   T
!
5370  gct  ttt  tct  agt  aat  tat  gat  tcc  ggt  gtt  tat  tct  tat  tta  acg
!
!      224 225 226 227 228 229 230 231 232 233 234 235 236 237 238
!      P   Y   L   S   H   G   R   Y   F   K   P   L   N   L   G
!
5415  cct  tat  tta  tca  cac  ggt  cgg  tat  ttc  aaa  cca  tta  aat  tta  ggt
!
!      239 240 241 242 243 244 245 246 247 248 249 250 251 252 253
35     Q   K   M   K   L   T   K   I   Y   L   K   K   F   S   R
!
5460  cag  aag  atg  aaa  tta  act  aaa  ata  tat  ttg  aaa  aag  ttt  tct  cgc
!
!      254 255 256 257 258 259 260 261 262 263 264 265 266 267 268
40     V   L   C   L   A   I   G   F   A   S   A   F   T   Y   S
!
5505  gtt  ctt  tgt  ctt  gcg  att  gga  ttt  gca  tca  gca  ttt  aca  tat  agt
!
!      269 270 271 272 273 274 275 276 277 278 279 280 281 282 283
45     Y   I   T   Q   P   K   P   E   V   K   K   V   V   S   Q
!
5550  tat  ata  acc  caa  cct  aag  ccg  gag  gtt  aaa  aag  gta  gtc  tct  cag
!
!      284 285 286 287 288 289 290 291 292 293 294 295 296 297 298
!      T   Y   D   F   D   K   F   T   I   D   S   S   Q   R   L
!
5595  acc  tat  gat  ttt  gat  aaa  ttc  act  att  gac  tct  tct  cag  cgt  ctt
!
!      299 300 301 302 303 304 305 306 307 308 309 310 311 312 313
!      N   L   S   Y   R   Y   V   F   K   D   S   K   G   K   L
!
5640  aat  cta  agc  tat  cgc  tat  gtt  ttc  aag  gat  tct  aag  gga  aaa  TTA
55                                     PacI
!
!      314 315 316 317 318 319 320 321 322 323 324 325 326 327 328
!      I   N   S   D   D   L   Q   K   Q   G   Y   S   L   T   Y

```

10045674-100501

```

5685 ATT Aat agc gac gat tta cag aag caa ggt tat tca ctc aca tat
!      PacI
!
!      329 330 331 332 333 334 335 336 337 338 339 340 341 342 343
5 !      i I  D  L  C  T  V  S  I  K  K  G  N  S  N  E
!      iv                                     M1 K
5730 att gat tta tgt act gtt tcc att aaa aaa ggt aat tca aAT Gaa
!                                     Start IV
!
!      344 345 346 347 348 349
10 !      i  I  V  K  C  N  .End of I
!      iv  L3 L  N5 V  I7 N  F  V10
5775 att gtt aaa tgt aat TAA T TTT GTT
!      IV continued.....
15 5800 ttc ttg atg ttt gtt tca tca tct tct ttt gct cag gta att gaa atg
5848 aat aat tcg cct ctg cgc gat ttt gta act tgg tat tca aag caa tca
5896 ggc gaa tcc gtt att gtt tct ccc gat gta aaa ggt act gtt act gta
5944 tat tca tct gac gtt aaa cct gaa aat cta cgc aat ttc ttt att tct
5992 gtt tta cgt gct aat aat ttt gat atg gtt ggt tca att cct tcc ata
20 6040 att cag aag tat aat cca aac aat cag gat tat att gat gaa ttg cca
6088 tca tct gat aat cag gaa tat gat gat aat tcc gct cct tct ggt ggt
6136 ttc ttt gtt cgg caa aat gat aat gtt act caa act ttt aaa att aat
6184 aac gtt cgg gca aag gat tta ata cga gtt gtc gaa ttg ttt gta aag
6232 tct aat act tct aaa tcc tca aat gta tta tct att gac ggc tct aat
25 6280 cta tta gtt gtt TCT gca cct aaa gat att tta gat aac ctt cct caa
!      ApaLI removed
6328 ttc ctt tct act gtt gat ttg cca act gac cag ata ttg att gag ggt
6376 ttg ata ttt gag gtt cag caa ggt gat gct tta gat ttt tca ttt gct
6424 gct ggc tct cag cgt ggc act gtt gca ggc ggt gtt aat act gac cgc
30 6472 ctc acc tct gtt tta tct tct gct ggt ggt tcg ttc ggt att ttt aat
6520 ggc gat gtt tta ggg cta tca gtt cgc gca tta aag act aat agc cat
6568 tca aaa ata ttg tct gtg cca cgt att ctt acg ctt tca ggt cag aag
6616 ggt tct atc tct gTt GGC CAG aat gtc cct ttt att act ggt cgt gtg
!      MscI
35 6664 act ggt gaa tct gcc aat gta aat aat cca ttt cag acg att gag cgt
6712 caa aat gta ggt att tcc atg agc gtt ttt cct gtt gca atg gct ggc
6760 ggt aat att gtt ctg gat att acc agc aag gcc gat agt ttg agt tct
6808 tct act cag gca agt gat gtt att act aat caa aga agt att gct aca
6856 acg gtt aat ttg cgt gat gga cag act ctt tta ctc ggt ggc ctc act
40 6904 gat tat aaa aac act tct caa gat tct ggc gta ccg ttc ctg tct aaa
6952 atc cct tta atc ggc ctc ctg ttt agc tcc cgc tct gat tcc aac gag
7000 gaa agc acg tta tac gtg ctc gtc aaa gca acc ata gta cgc gcc ctg
7048 TAG cggcgcatt
!      End IV
45 7060 aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccttagc
7120 gcccgctcct ttcgctttct tcccttcctt tctcgccacg ttcGCCGGCt ttccccgtca
!      NgoMI
7180 agctctaaat cgggggctcc cttaggggtt ccgatttagt gctttacggc acctcgaccc
7240 caaaaaactt gatttggttg atggttCACG TAGTGggcca tcgcctgat agacggtttt
50 !      DraIII
7300 tcgcccctttG ACGTTGGAGT Ccaggttctt taatagtgga ctcttggtcc aaactggaac
!      DrdI
7360 aacactcaac cctatctcgg gctattcttt tgatttataa gggattttgc cgatttcgga
7420 accaccatca aacaggattt tcgcctgctg gggcaaacca gcgtggaccg cttgctgcaa
55 7480 ctctctcagg gccaggcggt gaagggaact CAGCTGttgc cCGTCTCact ggtgaaaaga
!      PvuII.      BsmBI.
7540 aaaaccaccc tGGATCC AAGCTT
!      BamHI      HindIII (1/2)

```

1046740504

```

!               Insert carrying bla gene
7563   gcagggtg gcacttttcg gggaaatgtg cgcggaaccc
7600   ctatttgttt atttttctaa atacattcaa atatGTATCC gctcatgaga caataaccct
!               BciVI
5       7660   gataaatgct tcaataatat tgaaaaAGGA AGAgt
!               RBS.?...
!       Start bla gene
7695   ATG agt att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg gca ttt
7746   tgc ctt cct gtt ttt gct cac cca gaa acg ctg gtg aaa gta aaa gat gct
10      7797   gaa gat cag ttg ggC gCA CGA Gtg ggt tac atc gaa ctg gat ctc aac agc
!               BssSI...
!               ApaLI removed
7848   ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc
7899   act ttt aaa gtt ctg cta tgt cat aca cta tta tcc cgt att gac gcc ggg
15      7950   caa gaG CAA CTC GGT CGc cgg gcg cgg tat tct cag aat gac ttg gtt gAG
!               BcgI
!               ScaI
8001   TAC Tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga gaa
!               ScaI
8052   tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac tta ctt
20      8103   ctg aca aCG ATC Gga gga ccg aag gag cta acc gct ttt ttg cac aac atg
!               PvuI
8154   ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat gaa gcc
8205   ata cca aac gac gag cgt gac acc acg atg cct gta gca atg cca aca acg
8256   tTG CGC Aaa cta tta act ggc gaa cta ctt act cta gct tcc cgg caa caa
25      !       FspI....
!
8307   tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca ctt ctg cgc tcg
8358   GCC ctt ccG Gct ggc tgg ttt att gct gat aaa tct gga gcc ggt gag cgt
!               BglI
30      8409   gGG TCT Cgc ggt atc att gca gca ctg ggg cca gat ggt aag ccc tcc cgt
!               BsaI
8460   atc gta gtt atc tac acG ACg ggg aGT Cag gca act atg gat gaa cga aat
!               AhdI
35      8511   aga cag atc gct gag ata ggt gcc tca ctg att aag cat tgg TAA ctgt
!               stop
8560   cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt taatttaaaa
8620   ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatcccttaa cgtgagtttt
8680   cgttccactg tacgtaagac cccc
8704   AAGCTT GTCGAC tgaa tggcgaatgg cgctttgcct
40      !       HindIII SalI..
!       (2/2) HincII
8740   ggtttccggc accagaagcg gtgccggaaa gctggctgga gtgcgatctt
!
8790   CCTGAGG
45      !       Bsu36I
8797   ccgat actgtcgtcg tcccctcaaa ctggcagatg
8832   cacggttacg atgcgcccac ctacaccaac gtaacctatc ccattacggt caatccgccc
8892   tttgttccca cggagaatcc gacgggttgt tactcgctca catttaattgt tgatgaaagc
8952   tggctacagg aaggccagac gcgaattatt tttgatggcg ttccatttgg ttaaaaaatg
50      9012   agctgattta acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaATTAA
!               SwaI...
9072   Tatttgctta tacaatcttc ctgttttttg ggcttttctg attatcaacc GGGGTAc
!               RBS?
9131   ATG att gac atg cta gtt tta cga tta ccg ttc atc gat tct ctt gtt tgc
55      !       Start gene II
9182   tcc aga ctc tca ggc aat gac ctg ata gcc ttt gta GAT CTc tca aaa ata
!               BglII...
9233   gct acc ctc tcc ggc atg aat tta tca gct aga acg gtt gaa tat cat att

```

1004674-1004674



Table 21B: Sequence of MALIA3, condensed  
LOCUS MALIA3 9532

CIRCULAR

ORIGIN

1	AATGCTACTA	CTATTAGTAG	AATTGATGCC	ACCTTTTTCAG	CTCGCGCCCC	AAATGAAAAAT
5 61	ATAGCTAAAC	AGGTTATGTA	CCATTGCGA	AATGTATCTA	ATGGTCAAAC	TAAATCTACT
121	CGTTCGCAGA	ATTGGGAATC	AACTGTTACA	TGGAATGAAA	CTTCCAGACA	CCGTACTIONT
181	GTTGCATATT	TAAAACATGT	TGAGCTACAG	CACCAGATTC	AGCAATTAAG	CTCTAAGCCA
241	TCCGCAAAAA	TGACCTCTTA	TCAAAGGAG	CAATTAAAGG	TACTCTCTAA	TCCTGACCTG
361	TCTTTCGGGC	TTCTCTTAA	TCTTTTGTG	GCAATCCGCT	TTGCTTCTGA	CTATAATAGT
10 421	CAGGGTAAAG	ACCTGATTTT	TGATTTATGG	TCATTCTCGT	TTTCTGAACT	GTTTAAAGCA
481	TTTGAGGGGG	ATTCAATGAA	TATTTATGAC	GATTCCGCAG	TATTGGACGC	TATCCAGTCT
541	AAACATTTTA	CTATTACCCC	CTCTGGCAAA	ACTTCTTTTG	CAAAAGCCTC	TCGCTATTTT
601	GGTTTTTATC	GTCGTCTGGT	AAACGAGGGT	TATGATAGTG	TTGCTCTTAC	TATGCCTCGT
661	AATTCCTTTT	GGCGTTATGT	ATCTGCATTA	GTTGAATGTG	GTATTCCTAA	ATCTCAACTG
15 721	ATGAATCTTT	CTACCTGTAA	TAATGTTGTT	CCGTTAGTTC	GTTTTATTA	CGTAGATTTT
781	TCTTCCCAAC	GTCCTGACTG	GTATTAATGAG	CCAGTTCTTA	AAATCGCATA	AGGTAATTCA
841	CAATGATTAA	AGTTGAAATT	AAACCATCTC	AAGCCCAATT	TACTACTCGT	TCTGGTGTTC
901	CTCGTCAGGG	CAAGCCTTAT	TCATGAATG	AGCAGCTTTG	TTACGTTGAT	TTGGGTAATG
961	AATATCCGGT	TCTTGTCAAG	ATTACTCTTG	ATGAAGGTCA	GCCAGCCTAT	GCGCCTGGTC
20 1021	TGTACACCGT	TCATCTGTCC	TCTTTCAAAG	TTGGTCAGTT	CGGTTCCCTT	ATGATTGACC
1081	GTCTGCGCCT	CGTTCGGGCT	AAGTAACATG	GAGCAGGTG	CGGATTTCGA	CACAATTTAT
1141	CAGGCGATGA	TACAAATCTC	CGTTGTACTT	TGTTTCGCGC	TTGGTATAAT	CGCTGGGGGT
1201	CAAAGATGAG	TGTTTTAGTG	TATCTTTTCG	CCTCTTTCGT	TTTAGGTTGG	TGCCTTCGTA
1261	GTGGCATTAC	GTATTTTACC	CGTTAATGG	AAACTTCCTC	ATGAAAAAGT	CTTTAGTCTT
25 1321	CAAAGCCTCT	GTAGCCGTTG	CTACCCCTCG	TCCGATGCTG	TCTTTCGCTG	CTGAGGGTGA
1381	CGATCCCAGC	AAAGCGGCCT	TTAATCCCTT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTTA
1441	TGCGTGGGCG	ATGGTTGTTG	TCATTGTCCG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA
1501	ATTCACCTCG	AAAGCAAGCT	GATAAACCGA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTTT
1561	TTTTTGGAGA	TTTTCAACGT	GAAAAATTA	TTATTTCGCA	TTCTTTTAGT	TGTTCTTTTC
30 1621	TATTCTCACA	GTGCACAGTC	TGTCGTGACG	CAGCCGCCCT	CAGTGTCTGG	GGCCCCAGGG
1681	CAGAGGGTCA	CCATCTCCTG	CACTGGGAGC	AGCTCCAACA	TCGGGGCAGG	TTATGATGTA
1741	CACGTGGTACC	AGCAGCTTCC	AGGAACAGCC	CCCAAACTCC	TCATCTATGG	TAACAGCAAT
1801	CGGCCCTCAG	GGGTCCCTGA	CCGATTCTCT	GGCTCCAAGT	CTGGCACCTC	AGCCTCCCTG
1861	GCCATCACTG	GGCTCCAGGC	TGAGGATGAG	GCTGATTATT	ACTGCCAGTC	CTATGACAGC
35 1921	AGCCTGAGTG	GCCTTTTATG	CTTCGGAAC	GGGACCAAGG	TCACCCCTCT	AGGTCAGCCC
1981	AAGGCCAACC	CCACTGTCAC	TCTGTTCCCG	CCCTCCTCTG	AGGAGCTCCA	AGCCAACAAG
2041	GCCACACTAG	TGTGTCTGAT	CAGTGACTTC	TACCCGGGAG	CTGTGACAGT	GGCCTGGAAG
2101	GCAGATAGCA	GCCCCGTCAA	GGCGGGAGTG	GAGACCACCA	CACCCCTCAA	ACAAAGCAAC
2161	AACAAGTACG	CGGCCAGCAG	CTATCTGAGC	CTGACGCCTG	AGCAGTGGA	GTCCACAGA
40 2221	AGTACAGCT	GCCAGGTCAC	GCATGAAGGG	AGCACCGTGG	AGAAGACAGT	GGCCCCCTACA
2281	GAATGTTTAT	AATAAACCGC	CTCCACCGGG	CGCGCCAATT	CTATTTCAAG	GAGACAGTCA
2341	TAATGAAATA	CCTATTGCCT	ACGGCAGCCG	CTGGATTGTT	ATTACTCGCG	GCCCAGCCGG
2401	CCATGGCCGA	AGTTCAATTG	TTAGAGTCTG	GTGGCGGTCT	TGTTTCAGCT	GGTGGTTCTT
2461	TACGTCTTTC	TTGCGCTGCT	TCCGGATTCA	CTTTCTCTTC	GTACGCTATG	TCTTGGGTTT
45 2521	GCCAAGCTCC	TGGTAAAGGT	TTGGAGTGGG	TTTCTGCTAT	CTCTGTTTCT	GGTGGCAGTA
2581	CTTACTATGC	TGACTCCGTT	AAAGGTCGCT	TCATATCTC	TAGAGACAAC	TCTAAGAATA
2641	CTCTCTACTT	GCAGATGAAC	AGCTTAAGGG	CTGAGGACAC	TGCAGTCTAC	TATTGCGCTA
2701	AAGACTATGA	AGGTACTGGT	TATGCTTTTC	ACATATGGGG	TCAAGGTACT	ATGGTCACCG
2761	TCTCTAGTGC	CTCCACCAAG	GGCCCATCGG	TCTTCCCCCT	GGCACCTCTC	TCCAAGAGCA
50 2821	CCTCTGGGGG	CACAGCGGCC	CTGGGCTGCC	TGGTCAAGGA	CTACTTCCCC	GAACCGGTGA
2881	CGGTGTCTGT	GAATCAGGC	GCCCTGACCA	GCGGCGTCCA	CACCTTCCCC	GCTGTCTTAC
2941	AGTCTAGCGG	ACTCTACTCC	CTCAGCAGCG	TAGTGACCGT	GCCCTCTTCT	AGCTTGGGCA
3001	CCGAGACCTA	CATCTGCAAC	GTGAATCACA	AGCCCAGCAA	CACCAAGGTG	GACAAGAAAG
3061	TTGAGCCCAA	ATCTTGTGCG	GCCGCTCATC	ACCACCATCA	TCACTCTGCT	GAACAAAAAC
55 3121	TCATCTCAGA	AGAGGATCTG	AATGGTGCCG	CAGATATCAA	CGATGATCGT	ATGGCTGGCG
3181	CCGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	ACCCCATAC	AGAAAATTCA	TTTACTAACG
3241	TCTGGAAAGA	CGACAAAAC	TTAGATCGTT	ACGCTAACTA	TGAGGGTTGT	CTGTGGAATG

Table 21B: Sequence of MALIA3, condensed



5	3301	CTACAGGCGT	TGTAGTTTGT	ACTGGTGACG	AAACTCAGTG	TTACGGTACA	TGGGTTCTCA
	3361	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	TCTGAGGGTG
	3421	GCGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	ATTCGGGGCT
	3481	ATACCTATAT	CAACCTCTC	GACGGCACTT	ATCCGCCTGG	TACTAGCAA	AACCCCGCTA
	3541	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	CAGAATAATA
10	3601	GGTTCGAAA	TAGGCAGGGG	GCATTAACTG	TTTATACGGG	CACTGTTACT	CAAGGCACTG
	3661	ACCCCGTTAA	AACCTATTAC	CAGTACACTC	CTGTATCATC	AAAAGCCATG	TATGACGCTT
	3721	ACTGGAACGG	TAAATTTCAGA	GACTGCGCTT	TCCATTCTGG	CTTTAATGAA	GATCCATTCTG
	3781	TTTGTGAATA	TCAAGGCCAA	TCGTCTGACC	TGCCTCAACC	TCCTGTCAAT	GCTGGCGGGC
	3841	GCTCTGGTGG	TGGTCTGGT	GGCGGCTCTG	AGGCTGGTGG	CTCTGAGGGT	GGCGGTTCTG
15	3901	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTTCGGGT	GATTTTGTAT
	3961	ATGAAAAGAT	GGCAAACGCT	AATAAGGGGG	CTATGACCGA	AAATGCCGAT	GAAAACGCGC
	4021	TACAGTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTCGCTAC	TGATTACGGT	GCTGCTATCG
	4081	ATGGTTTCAT	TGGTGACGTT	TCCGCGCTTG	CTAATGGTAA	TGGTGCTACT	GGTGATTTTG
	4141	CTGGCTCTAA	TCCCAAAATG	GCTCAAGTCG	GTGACGGTGA	TAATTACACT	TTAATGAATA
20	4201	ATTTCCGTCA	ATATTTACCT	TCCCTCCCTC	AATCGGTTGA	ATGTCGCCCT	TTTGCTCTTA
	4261	GCGCTGGTAA	ACCATATGAA	TTTTCTATTG	ATTGTGACAA	AATAAACTTA	TTCCGTGGTG
	4321	TCTTTGCGTT	TCCTTTATAT	GTTGCCACCT	TTATGTATGT	ATTTTCTACG	TTTGCTAACA
	4381	TACTGCGTAA	TAAGGAGTCT	TAATCATGCC	AGTTCTTTTG	GGTATTCGGT	TATTATTGCG
	4441	TTTCCCTCGT	TTCCCTCTGG	TAACTTTGTT	CGGCTATCTG	CTTACTTTTC	TTAAAAGGG
25	4501	CTTCCGTAA	ATAGCTATTG	CTATTTCATT	GTTTCTTGCT	CTTATTATTG	GGCTTAACTC
	4561	AATTCTTGTG	GGTTATCTCT	CTGATATTAG	CGCTCAATTA	CCCTCTGACT	TTGTTCAGGG
	4621	TGTTCAGTTA	ATTCTCCCGT	CTAATGCGCT	TCCTGTTTTT	TATGTTATTC	TCTCTGTA
	4681	GGCTGCTATT	TTCATTTTTT	ACGTTAAACA	AAAAATCGTT	TCTTATTTGG	ATTGGGATAA
	4741	ATAATATGGC	TGTTTATTTT	GTAACGTGCA	AATTAGGCTC	TGGAAGAGC	CTCGTTAGCG
30	4801	TTGGTAAAGT	TACAGATAAA	ATTATGAGTG	GGTGCAAAAT	AGCAACTAAT	CTTGATTTAA
	4861	GGCTTCAAAA	CCTCCCGCAA	GTCGGGAGGT	TCGCTAAAC	GCCTCGCGTT	CTTAGAATAC
	4921	CGGATAAGCC	TTCTATATCT	GATTTGCTTG	CTATTGGGCG	CGGTAATGAT	TCCTACGATG
	4981	AAAATAAAAA	CGGCTTGCTT	GTTCTCGATG	AGTGCGGTAC	TTGGTTTAA	ACCGTTCTT
	5041	GGAATGATAA	GGAAAGACAG	CCGATTATTG	ATTGGTTTCT	ACATGCTCGT	AAATTAGGAT
35	5101	GGGATATTAT	TTTTCTTGTT	CAGGACTTAT	CTATTGTTCT	TAAACAGCGT	CGTCTGCAT
	5161	TAGCTGAACA	TTTGTTTTAT	TGTCGTCGTC	TGGACGAAT	TACTTTACCT	TTTGTCGGTA
	5221	CTTTATATTC	TCTTATTACT	GGCTCGAAAA	TGCCTCTGCC	TAAATTACAT	GTTGGCGTTG
	5281	TTAAATATGG	CGATTCTCAA	TTAAGCCCTA	CTGTTGAGCG	TTGGCTTTAT	ACTGGTAAGA
	5341	ATTTGTATAA	CGCATATGAT	ACTAAACAGG	CTTTTCTAG	TAATTATGAT	TCCGGTGTTT
40	5401	ATTTCTATTT	AACGCCTTAT	TTATCACACG	GTCGGTATTT	CAAACCATTA	AATTTAGGTT
	5461	AGAAGATGAA	ATTAACATAA	ATATATTTGA	AAAAGTTTTT	TCGCGTTCTT	TGTTTGC
	5521	TTGGATTTGC	ATCAGCATTT	ACATATAGTT	ATATAACCCA	ACCTAAGCCG	GAGGTTAAAA
	5581	AGGTAGTCTC	TCAGACCTAT	GATTTTGATA	AATTCACAT	TGACTCTTCT	CAGCGTCTTA
	5641	ATCTAAGCTA	TCGCTATGTT	TTCAAGGATT	CTAAGGGAAA	ATTAATTAAT	AGCGACGATT
45	5701	TACAGAAGCA	AGGTTATTCA	CTCATATATA	TGATTTTATG	TACTGTTTCC	ATTA
	5761	GTAATTCAAA	TGAAATTTGT	AAATGTAATT	AATTTTGTTT	TCTTGATTGT	TGTTCTATCA
	5821	TCTTCTTTTG	CTCAGGTAAT	TGAAATGAAT	AATTGCGCTC	TGCGCGATTT	TGTA
	5881	TATTC	AATCAGGCGA	ATCCGTTATT	GTTTCTCCCG	ATGTAAAGG	TACTGTTACT
	5941	GTATATTCAT	CTGACGTTAA	ACCTGAA	CTACGCAATT	TCTTTATTTT	TGTTTACGT
50	6001	GCTAAAT	TTGATATGGT	TGGTTCAATT	CCTTCCATAA	TTCAGAAGTA	TAATCCAAAC
	6061	AATCAGGATT	ATATTGATGA	ATTGCCATCA	TCTGATAATA	AGGAATATGA	TGATAATTC
	6121	GCTCCTTCTG	GTGGTTTCTT	TGTTCCGCAA	AATGATAATG	TTACTCAAAC	TTTTAAAT
	6181	AATAACGTTT	GGGCAAAGGA	TTTAATACGA	GTTGTCGAAT	TGTTTGTA	GTCTAATACT
	6241	TCTAAATCCT	CAAATGTATT	ATCTATTGAC	GGCTCTAATC	TATTAGTTGT	TTCTGCACCT
55	6301	AAAGATATTT	TAGATAACCT	TCCTCAATT	CTTTCTACTG	TTGATTTTGC	AAC
	6361	ATATTGATTG	AGGGTTTGAT	ATTTGAGGTT	CAGCAAGGTT	ATGCTTTAGA	TTTTCTATT
	6421	GCTGCTGGCT	CTCAGCGTGG	CAC	GCGGTTGTTA	ATACTGACCG	CCTCACCTCT
	6481	GTTTTATCTT	CTGCTGGTGG	TTCGTTCCGT	ATTTTAAATG	GCGATGTTTT	AGGGCTATCA
	6541	GTTTCGCGCAT	TAAAGACTAA	TAGCCATTCA	AAAATATTGT	CTGTGCCACG	TAT

6781	ACCAGCAAGG	CCGATAGTTT	GAGTCTTCT	ACTCAGGCAA	GTGATGTTAT	TACTAATCAA
6841	AGAAGTATTG	CTACAACGGT	TAATTTGCGT	GATGGACAGA	CTCTTTTACT	CGGTGGCCTC
6901	ACTGATTATA	AAAACACTTC	TCAAGATTCT	GGCGTACCGT	TCCTGTCTAA	AATCCCTTTA
6961	ATCGGCCCTC	TGTTTAGCTC	CCGCTCTGAT	TCCAACGAGG	AAAGCACGTT	ATACGTGCTC
5 7021	GTCAAAGCAA	CCATAGTACG	CGCCCTGTAG	CGGCGCATT	AGCGCGGCGG	GTGTGGTGGT
7081	TACGCGCAGC	GTGACCGCTA	CACCTGCCAG	CGCCCTAGCG	CCCCTCCTT	TCGCTTTCTT
7141	CCCTTCCTTT	CTCGCCACGT	TCGCCGGCTT	TCCCCGTCAA	GCTCTAAATC	GGGGGCTCCC
7201	TTTAGGGTTC	CGATTTAGTG	CTTTACGGCA	CCTCGACCCC	AAAAAACTTG	ATTTGGGTGA
7261	TGGTTCACGT	AGTGGGCCAT	CGCCCTGATA	GACGGTTTTT	CGCCCTTTGA	CGTTGGAGTC
10 7321	CACGTTCTTT	AATAGTGGAC	TCTTGTTC	AACTGGAACA	ACACTCAACC	CTATCTCGGG
7381	CTATTTCTTT	GATTTATAAG	GGATTTTGCC	GATTTGCGAA	CCACCATCAA	ACAGGATTTT
7441	CGCCTGCTGG	GGCAAACCAG	CGTGGACCGC	TTGCTGCAAC	TCTCTCAGGG	CCAGGCGGTG
7501	AAGGGCAATC	AGCTGTTGCC	CGTCTCACTG	GTGAAAAGAA	AAACCACCC	GGATCCAAGC
7561	TTGCAGGTGG	CACTTTTTCG	GGAAATGTGC	GCGGAACCCC	TATTTGTTTA	TTTTTCTAAA
15 7621	TACATTCAAA	TATGTATCCG	CTCATGAGAC	AATAACCCCTG	ATAAATGCTT	CAATAATATT
7681	GAAAAAGGAA	GAGTATGAGT	ATTCAACATT	TCCGTGTGCG	CCTTATTTCC	TTTTTTGCGG
7741	CATTTTGCCT	TCCTGTTTTT	GCTCACCCAG	AAACGCTGGT	GAAAGTAAAA	GATGCTGAAG
7801	ATCAGTTGGG	CGCACGAGTG	GGTTACATCG	AACTGGATCT	CAACAGCGGT	AAGATCCTTG
7861	AGAGTTTTTC	CCCCGAAGAA	CGTTTTTCAA	TGATGAGCAC	TTTTAAAGTT	CTGCTATGTC
20 7921	ATACACTATT	ATCCCGTATT	GACGCCGGGC	AAGAGCAACT	CGGTCGCCGG	GCGCGGTATT
7981	CTCAGAATGA	CTTGTTGAG	TACTCACCAG	TCACAGAAAA	GCATCTTACG	GATGGCATGA
8041	CAGTAAGAGA	ATTATGCAGT	GCTGCCATAA	CCATGAGTGA	TAACACTGCG	GCCAACTTAC
8101	TTCTGACAAC	GATCGGAGGA	CCGAAGGAGC	TAACCGCTTT	TTTGCACAAC	ATGGGGGATC
8161	ATGTAACCTG	CCTTGATCGT	TGGGAACCGG	AGCTGAATGA	AGCCATACCA	AACGACGAGC
25 8221	GTGACACCAC	GATGCCTGTA	GCAATGCCAA	CAACGTTGCG	CAAACATATTA	ACTGGCGAAC
8281	TACTTACTCT	AGCTTCCCGG	CAACAATTAA	TAGACTGGAT	GGAGGCGGAT	AAAGTTGCAG
8341	GACCACTTCT	GCGCTCGGCC	CTTCCGGCTG	GCTGGTTTAT	TGCTGATAAA	TCTGGAGCCG
8401	GTGAGCGTGG	GTCTCGCGGT	ATCATTGACG	CACTGGGGCC	AGATGGTAAG	CCCTCCCGTA
8461	TCGTAGTTAT	CTACACGACG	GGGAGTCAGG	CAACTATGGA	TGAACGAAAT	AGACAGATCG
30 8521	CTGAGATAGG	TGCCTCACTG	ATTAAGCATT	GGTAACGTGC	AGACCAAGTT	TACTCATATA
8581	TACTTTAGAT	TGATTTAAAA	CTTCATTTT	AATTTAAAA	GATCTAGGTG	AAGATCCTTT
8641	TTGATAATCT	CATGACCAAA	ATCCCTTAAC	GTGAGTTTTT	GTTCCACTGT	ACGTAAGACC
8701	CCCAAGCTTG	TCGACTGAAT	GGCGAATGGC	GCTTTGCCTG	GTTTCCGGCA	CCAGAAGCGG
8761	TGCCGAAAG	CTGGCTGGAG	TGCGATCTTC	CTGAGGCCGA	TACTGTCGTC	GTCCCTCAA
35 8821	ACTGGCAGAT	GCACGGTTAC	GATGCGCCCA	TCTACACCAA	CGTAACCTAT	CCCATTACGG
8881	TCAATCCGCC	GTTTGTTC	ACGGAGAATC	CGACGGGTG	TTACTCGCTC	ACATTTAATG
8941	TTGATGAAAG	CTGGCTACAG	GAAGGCCAGA	CGCGAATTAT	TTTTGATGGC	GTTCTATTG
9001	GTTAAAAAAT	GAGCTGATTT	AACAAAAATT	TAACGCGAAT	TTTAACAAAA	TATTAACGTT
9061	TACAATTTAA	ATATTTGCTT	ATACAATCTT	CCTGTTTTTG	GGGCTTTTCT	GATTATCAAC
40 9121	CGGGGTACAT	ATGATTGACA	TGCTAGTTTT	ACGATTACCG	TTCATCGATT	CTCTGTTTG
9181	CTCCAGACTC	TCAGGCAATG	ACCTGATAGC	CTTTGTAGAT	CTCTCAAAAA	TAGCTACCC
9241	CTCCGGCATG	AATTTATCAG	CTAGAACGGT	TGAATATCAT	ATTGATGGTG	ATTTGACTGT
9301	CTCCGGCCTT	TCTCACCCCT	TTGAATCTTT	ACCTACACAT	TACTCAGGCA	TTGCATTTAA
9361	AATATATGAG	GGTTCTAAAA	ATTTTATCC	TTGCGTTGAA	ATAAAGGCTT	CTCCCGCAAA
45 9421	AGTATTACAG	GGTCATAATG	TTTTTGGTAC	AACCGATTTA	GCTTTATGCT	CTGAGGCTTT
9481	ATTGCTTAAT	TTTGCTAATT	CTTGCCTTG	CCTGTATGAT	TTATTGGATG	TT

Table 22: Primers used in RACE amplification:

Heavy chain	
HuCh-FOR (1st PCR)	5'-TGG AAG AGG CAC GTT CTT TTC TTT-3'
HuCh-Nested (2nd PCR)	5' CTT TTC TTT GTT GCC GTT GGG GTG-3'
Kappa light chain	
HuKFor (1st PCR)	5'-ACA CTC TCC CCT GTT GAA GCT CTT-3'
HuKForAscI(2nd PCR)	5'-ACC GCC TCC ACC GGG CGC GGC TTA TTA ACA CTC TCC CCT GTT GAA GCT CTT-3'
Lambda light chain	
HuCLambdaFor (1st PCR)	5'-TGA ACA TTC TGT AGG GGC CAC TG-3'
HuCL2-FOR	5'-AGA GCA TTC TGC AGG GGC CAC TG-3'
HuCL7-FOR	
HuCLambdaForAscI (2nd PCR)	
HuCL2-FOR-ASC	5'-ACC GCC TCC ACC GGG CGC GCC TTA TTA TGA ACA TTC TGT AGG GGC CAC TG-3'
HuCL7-FOR-ASC	5'-ACC GCC TCC ACC GGG CGC GCC TTA TTA AGA GCA TTC TGC AGG GGC CAC TG-3'
GeneRacer 5' Primers provided with the kit (Invitrogen)	
5'A 1st PCR	5'CGACTGGAGCACGAGCACACTGA 3'
5'NA 2nd PCR	5'GGACACTGACATGGACTGAAGGAGTA-3'

Table 23: ONs used in Capture of kappa light chains using CJ method and *BsmAI*

All ONs are written 5' to 3'.

REdapters (6)  
 ON\_20SK15O12 gggAggATggAgAcTgggTc  
 ON\_20SK15L12 gggAAGATggAgAcTgggTc  
 ON\_20SK15A17 gggAgAgTggAgAcTgAgTc  
 ON\_20SK15A27 gggTgccTggAgAcTgcgTc  
 ON\_20SK15A11 gggTggcTggAgAcTgcgTc  
 ON\_20SK15BgggAgTcTggAgAcTgggTc

10 Bridges (6)  
 kapbr11O12 gggAggATggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg  
 kapbr11L12 gggAAGATggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg  
 kapbr11A17 gggAgAgTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg  
 kapbr11A27 gggTgccTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg  
 kapbr11A11 gggTggcTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg  
 kapbr11B3 gggAgTcTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg

Extender (5' biotinylated)  
 kapext1bio ccTcTgTcAcAgTgcAcAAgAcATccAgATgAcccAgTcTcc

20 Primers  
 kapPCRT1 ccTcTgTcAcAgTgcAcAAgAc 5'-aca ctc tcc cct gtt gaa gct ctt-3'  
 kapfor



Table 25: h3401-h2 captured Via CJ with BsmAI

! 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
! S A Q D I Q M T Q S P A T L S  
aGT GCA Caa gac atc cag atg acc cag tct cca gcc acc ctg tct  
5 ! ApaLI... a gcc acc ! L25,L6,L20,L2,L16,A11  
! Extender.....Bridge...  
  
! 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30  
! V S P G E R A T L S C R A S Q  
10 gtg tct cca ggg gaa agg gcc acc ctc tcc tgc agg gcc agt cag  
  
! 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45  
! S V S N N L A W Y Q Q K P G Q  
agt gtt agt aac aac tta gcc tgg tac cag cag aaa cct ggc cag  
15  
! 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60  
! V P R L L I Y G A S T R A T D  
gtt ccc agg ctc ctc atc tat ggt gca tcc acc agg gcc act gat  
  
20 ! 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75  
! I P A R F S G S G S G T D F T  
atc cca gcc agg ttc agt ggc agt ggg tct ggg aca gac ttc act  
  
! 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90  
25 ! L T I S R L E P E D F A V Y Y  
ctc acc atc agc aga ctg gag cct gaa gat ttt gca gtg tat tac  
  
! 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105  
! C Q R Y G S S P G W T F G Q G  
30 tgt cag cgg tat ggt agc tca cgg ggg tgg acg ttc ggc caa ggg  
  
! 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120  
! T K V E I K R T V A A P S V F  
acc aag gtg gaa atc aaa cga act gtg gct gca cca tct gtc ttc  
35  
! 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135  
! I F P P S D E Q L K S G T A S  
atc ttc ccg cca tct gat gag cag ttg aaa tct gga act gcc tct  
  
40 ! 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150  
! V V C L L N N F Y P R E A K V  
gtt gtg tgc ctg ctg aat aac ttc tat ccc aga gag gcc aaa gta  
  
! 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165  
45 ! Q W K V D N A L Q S G N S Q E  
cag tgg aag gtg gat aac gcc ctc caa tgg ggt aac tcc cag gag  
  
! 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180  
! S V T E Q D S K D S T Y S L S  
50 agt gtc aca gag cag gac agc aag gac agc acc tac agc ctc agc  
  
! 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195

Table 25: h3401-h2



Table 26: h3401-d8 KAPPA captured with CJ and *Bsm*AI

! 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
! S A Q D I Q M T Q S P A T L S  
5 aGT GCA Caa gac atc cag atg acc cag tct cct gcc acc ctg tct  
! ApaLI...Extender.....a gcc acc ! L25,L6,L20,L2,L16,A11  
! A GCC ACC CTG TCT ! L2

! 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30  
10 ! V S P G E R A T L S C R A S Q  
gtg tct cca ggt gaa aga gcc acc ctc tcc tgc agg gcc agt cag  
! GTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC ! L2

! 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45  
15 ! N L L S N L A W Y Q Q K P G Q  
aat ctt ctc agc aac tta gcc tgg tac cag cag aaa cct ggc cag

! 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60  
20 ! A P R L L I Y G A S T G A I G  
gct ccc agg ctc ctc atc tat ggt gct tcc acc ggg gcc att ggt

! 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75  
! I P A R F S G S G S G T E F T  
25 atc cca gcc agg ttc agt ggc agt ggg tct ggg aca gag ttc act

! 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90  
! L T I S S L Q S E D F A V Y F  
ctc acc atc agc agc ctg cag tct gaa gat ttt gca gtg tat ttc

! 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105  
30 ! C Q Q Y G T S P P T F G G G T  
tgt cag cag tat ggt acc tca ccg ccc act ttc ggc gga ggg acc

! 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120  
35 ! K V E I K R T V A A P S V F I  
aag gtg gag atc aaa cga act gtg gct gca cca tct gtc ttc atc

! 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135  
! F P P S D E Q L K S G T A S V  
40 ttc ccg cca tct gat gag cag ttg aaa tct gga act gcc tct gtt

! 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150  
! V C P L N N F Y P R E A K V Q  
gtg tgc ccg ctg aat aac ttc tat ccc aga gag gcc aaa gta cag  
45

! 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165  
! W K V D N A L Q S G N S Q E S  
tgg aag gtg gat aac gcc ctc caa tgc ggt aac tcc cag gag agt

! 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180  
50 ! V T E Q D N K D S T Y S L S S  
gtc aca gag cag gac aac aag gac agc acc tac agc ctc agc agc



! 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195  
! T L T L S K V D Y E K H E V Y  
acc ctg acg ctg agc aaa gta gac tac gag aaa cac gaa gtc tac  
5  
! 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210  
! A C E V T H Q G L S S P V T K  
gcc tgc gaa gtc acc cat cag ggc ctt agc tcg ccc gtc acg aag  
10  
! 211 212 213 214 215 216 217 218 219 220 221 222 223  
! S F N R G E C K K E F V  
agc ttc aac agg gga gag tgt aag aaa gaa ttc gtt t

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 27. V3-23 VH framework with variegated codons shown

```

!
!           17 18 19 20 21 22
!           A Q P A M A
5 5'-ctg tct gaa cG GCC cag ccG GCC atg gcc 29
   3'-gac aga ctt gc cgg gtc ggc cgg tac cgg
!       Scab.....SfiI.....
!           NgoMI...
!           NcoI....
10
!           FR1(DP47/V3-23)-----
!           23 24 25 26 27 28 29 30
!           E V Q L L E S G
!           gaa|gtt|CAA|TTG|tta|gag|tct|ggt| 53
15 ctt|caa|gtt|aac|aat|ctc|aga|cca|
   | MfeI |
!
! -----FR1-----
! 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
20 G G L V Q P G G S L R L S C A
! |ggc|ggt|ctt|ggt|cag|cct|ggt|ggt|tct|tta|cgt|ctt|tgc|gct| 98
! |cgg|cca|gaa|caa|gtc|gga|cca|cca|aga|aat|gca|gaa|aga|acg|cga|
!
! Sites to be varied--> *** *** ***
25 -----FR1----->|...CDR1.....|-----FR2-----
! 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
! A S G F T F S S Y A M S W V R
! |gct|TCC|GGA|ttc|act|ttc|tct|tCG|TAC|Gct|atg|tct|tgg|ggt|cgC| 143
! |cga|agg|cct|aag|tga|aag|aga|agc|atg|cga|tac|aga|acc|caa|gcg|
30 | BspEI | | BsiWI | | BstXI.
!
! Sites to be varies--> *** *** ***
! -----FR2----->|...CDR2.....
35 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
! Q A P G K G L E W V S A I S G
! |CAA|gct|ccT|GGT|aaa|ggt|ttg|gag|tgg|ggt|tct|gct|atc|tct|ggt| 188
! |ggt|cga|gga|cca|ttt|cca|aac|ctc|acc|caa|aga|cga|tag|aga|cca|
! ...BstXI |
!
! *** ***
40 .....CDR2.....|-----FR3-----
! 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
! S G G S T Y Y A D S V K G R F
! |tct|ggt|ggc|agt|act|tac|tat|gct|gac|tcc|ggt|aaa|ggt|cgc|ttc| 233
45 |aga|cca|cgg|tca|tga|atg|ata|cga|ctg|agg|caa|ttt|cca|gcg|aag|
!
! -----FR3-----
! 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
50 T I S R D N S K N T L Y L Q M
! |act|atc|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|tgc|cag|atg| 278
! |tga|tag|aga|tct|ctg|ttg|aga|ttc|tta|tga|gag|atg|aac|gtc|tac|
! | XbaI |
!
! -----FR3----->|
55 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
! N S L R A E D T A V Y Y C A K
! |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgc|gct|aaa| 323
! |ttg|tcg|aat|tcc|cga|ctc|ctg|tga|cgt|cag|atg|ata|acg|cga|ttt|

```

Table 27. V3-23 VH framework with variegated codons shown

```

!      |AflII|      |PstI|
!
!      ..... CDR3.....|-----FR4-----
!      121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
5 !      D Y E G T G Y A F D I W G Q G
!      |gac|tat|gaa|ggg|act|ggg|tat|gct|ttc|gaC|ATA|TGg|ggg|caa|ggg| 368
!      |ctg|ata|ctt|cca|tga|cca|ata|cga|aag|ctg|tat|acc|cca|ggt|cca|
!      |NdeI|
!
10 !      -----FR4----->|
!      136 137 138 139 140 141 142
!      T M V T V S S
!      |act|atG|GTC|ACC|gtc|tct|agt- 389
!      |tga|tac|cag|tgg|cag|aga|tca-
15 !      |BstEII|
!
!      143 144 145 146 147 148 149 150 151 152
!      A S T K G P S V F P
!      gcc tcc acc aaG GGC CcA tcg GTC TTC ccc-3' 419
20 !      cgg agg tgg ttc ccg ggt agc cag aag ggg-5'
!      Bsp120I. BbsI...(2/2)
!      ApaI....

(SFPRMET) 5'-ctg tct gaa cG GCC cag ccG-3'
(TOPFR1A) 5'-ctg tct gaa cG GCC cag ccG GCC atg gcc-
25 !      gaa|gtt|CAA|TTG|tta|gag|tct|ggg|-
!      |ggc|ggg|ctt|gtt|cag|cct|ggg|ggg|tct|tta-3'
(BOTFR1B) 3'-caa|gtc|gga|cca|cca|aga|aat|gca|gaa|aga|acg|cga|-
!      |cga|agg|cct|aag|tga|aag-5' ! bottom strand
(BOTFR2) 3'-acc|caa|gcg|-
30 !      |gtt|cga|gga|cca|ttt|cca|aac|ctc|acc|caa|aga|-5' ! bottom strand
(BOTFR3) 3'- a|cga|ctg|agg|caa|ttt|cca|gca|aag|-
!      |tga|tag|aga|tct|ctg|ttg|aga|ttc|tta|tga|gag|atg|aac|gtc|tac|-
!      |ttg|tcg|aat|tcc|cga|ctc|ctg|tga-5'
(F06) 5'-gC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgc|gct|aaa|-
35 !      |gac|tat|gaa|ggg|act|ggg|tat|gct|ttc|gaC|ATA|TGg|ggg|c-3'
(BOTFR4) 3'-cga|aag|ctg|tat|acc|cca|ggt|cca|-
!      |tga|tac|cag|tgg|cag|aga|tca-
!      cgg agg tgg ttc ccg ggt agc cag aag ggg-5' ! bottom strand
(BOTPRCPRIM) 3'-gg ttc ccg ggt agc cag aag ggg-5'
40 !
! CDR1 diversity
!
! (ON-vgC1) 5'-gct|TCC|GGA|ttc|act|ttc|tct|<1>|TAC|<1>|atg|<1>|-
!      CDR1.....6859
45 !      |tgg|gtt|cgC|CAa|gct|ccT|GG-3'
!
! <1> stands for an equimolar mix of {ADEF GHIKLMNPQRSTVWY}; no C
! (this is not a sequence)
!
50 ! CDR2 diversity
!
! (ON-vgC2) 5'-ggg|ttg|gag|tgg|gtt|tct|<2>|atc|<2>|<3>|-
!      CDR2.....
!      |tct|ggg|ggc|<1>|act|<1>|tat|gct|gac|tcc|gtt|aaa|gg-3'
55 !      CDR2.....
! <1> is an equimolar mixture of {ADEF GHIKLMNPQRSTVWY}; no C
! <2> is an equimolar mixture of {YRWVGS}; no ACDEFHIKLMNPQT

```

10463476666

! <3> is an equimolar mixture of {PS}; no ACDEFGHIKLMNQRTVWY

Table 28: Stuffer used in VH

1 TCCGGAGCTT CAGATCTGTT TGCCTTTTGT TGGGGTGGTG CAGATCGCGT TACGGAGATC  
61 GACCGACTGC TTGAGCAAAA GCCACGCTTA ACTGCTGATC AGGCAI GGGA TGTTATTCGC  
121 CAAACCAAGTC GTCAGGATCT TAACTGAGG CTTTTTTTAC CTACTCTGCA AGCAGGACA  
181 TCTGGTTTGA CACAGAGCGA TCCGGTCTGT CAGTTGGTAG AACATTAAAC ACGTTGGGAT  
241 GGCAATCAATT TGCTTAATGA TGATGGTAAA ACCTGGCAGC AGCCAGGCTC TGCCATCCTG  
301 AACGTTTGGC TGACCAAGTAT GTTGAAGCGT ACCGTAGTGG CTGCCGTACC TATGCCATTT  
361 GATAAGTGGT ACAGCGCCAG TGGCTACGAA ACAACCCAGG ACGCCCCAAC TGGTTCGCTG  
421 AATATAAGTG TTGGAGCAAA AATTTGTAT GAGCGGTGC AGGAGACAA ATCACCAATC  
481 CCACAGGCGG TTGATCTGTT TGCTGGGAAA CCACAGCAGG AGGTTGTGTT GGCTGCGCTG  
541 GAAGATACCT GGGAGACTCT TTCCAACGC TATGGCAATA ATGTAGTAA CTGGAATAA  
601 CCTGCAATGG CCTTAACGTT CCGGGCAAT AATTCTTTG GTGTACCGCA GGCCGCGACG  
661 GAAGAAACGC GTCATCAGG GGAGTATCAA AACCGTGAA CAGAAAACGA TATGATTGTT  
721 TTCTCACCAA CGACAAGCGA TCGTCTGTG CTTCCTGGG ATGTGGTCGC ACCCGGTCAG  
781 AGTGGGTTTA TTGCTCCCGA TGGAAACAGT GATAAGCACT ATGAAGATCA GCTGAAAATG  
841 TACGAAAATT TTGGCCGTAA GTCGCTCTGG TTAACGAAAG AGGATGTGGA GCGGCATAAG  
901 GAGTCGTCTA GA

5

10

15

Table 29: DNA sequence of pCES5

```

! pCES5 6680 bases = pCes4 with stuffers in CDR1-2 and CDR3 2000.12.13
!
! Ngene = 6680
! Useful REs (cut MAAnoLI fewer than 3 times) 2000.06.05
!
! Non-cutters
! Acc65I Ggtacc Afel AGCgct AvriI Cctagg
! BsaBI GATNNnmtc BsaWI Cgtacg BsmFI Nnnnnnnnnnnnnngtccc
! BsrGI Tgtaca BstAPI GCANNNNNntgc BstBI TTegaa
! BstZ17I GTAtac BrlI CACgtg Ecl136I GAGtc
! EcoRV GATatc FseI GGCCGGcc KpnI GGTAAC
! MscI TGGcca NruI TCGcca NsiI ATGCAI
! PacI TTAAThaa PmeI GTTaaac PmlI CACgtg
! PpuMI RCGwccy PshAI GACNNnnngtc SacI GAGCTc
! SaeII CCGCgg SbfI CCTGCagg SexAI Accwgt
! SgfI GCGATgc SnaBI TACgta SpeI Actagt
! SphI GCATGc Sse8387I CCTGCagg StuI AGGcct
! SwaI ATTaaat XmaI Ccggg
!
! cutters
! Enzymes that cut more than 3 times
! AlwNI CAGNNNctg 5
! BsgI ctgcac 4
! BsrFI Rccggy 5
! EarI CTCCTTCNnn 4
! PfuI nNNNNNNGCGGG 10
!
! Enzymes that cut from 1 to 3 times.
!
! EcoO109I RGnccy 3 7 2636 4208
! BssSI Ctegtg 1 12
! "- Cagag 1 1703
! BspHI Tcatga 3 43 148 1156

```

5	!Aatl GACGtc	1	65
	!Bcvi GTATCCNNNNNN	2	140 1667
	!Eco57t CTGAAG	1	301
	!"- cttag	2	1349
	!Aval Cyegrg	3	319 2347 6137
	!BsiHKAi GWGCWc	3	401 2321 4245
	!HgiAI GWGCWc	3	401 2321 4245
	!Bcgl gcanunnnteg	1	461
	!Scal AGTact	1	505
10	!PvuI CGATeg	3	616 3598 5926
	!Fspl TGCgca	2	763 5946
	!BglII GCCNNNNnggc	3	864 2771 5952
	!Bpml CTGGAG	1	898
	!"- cttag	1	4413
15	!Bsal GGTCTCNnnnn	1	916
	!AhdI GACNNNngtc	1	983
	!Eam1105I GACNNNngtc	1	983
	!DrdI GACNNNNngtc	3	1768 6197 6579
	!SapI gaagagc	1	1998
20	!PvuI CAGctg	3	2054 3689 5896
	!PmlI CCANNNNttg	3	2233 3943 3991
	!HindIII Aagctt	1	2235
	!Apa1I Gtcgac	1	2321
	!BspMI Nnnnnnnngcaggt	1	2328
	!"- ACCTGCNNNNn	2	3460
25	!PstI CTGCag	1	2335
	!AclI GTmkac	2	2341 2611
	!HmcII GTYrac	2	2341 3730
	!Sall Gtcgac	1	2341
30	!TliI Ctcgag	1	2347
	!XhoI Ctcgag	1	2347
	!BbsI gcttc	2	2383 4219
	!BlpI GCtnagc	1	2580
	!EspI GCtnagc	1	2580
35	!SgrAI CRccggyg	1	2648
	!AclI Accggt	2	2649 4302

1 Ascl GGcgcc 1 2689  
 1 BssHII Gcgcc 1 2690  
 1 SfiI GGCCNNNNnggcc 1 2770  
 2 2776 6349  
 1 NaeI GCCggc 2 2776 6349  
 1 NgoMIV Gcggc 2 2776 6349  
 1 BglI Corygg 3 2781 3553 5712  
 1 Dsal Ccrgg 3 2781 3553 5712  
 1 NcoI Ccatgg 1 2781  
 1 StyI Ccwvvg 3 2781 4205 4472  
 1 MfeI Caatg 1 2795  
 1 BspEI Tcggga 1 2861  
 1 BglII Agatct 1 2872  
 1 BclI Tgatca 1 2956  
 1 Bsu36I CCtnagg 3 3004 4143 4373  
 1 XcmI CCANNNNNnnntgg 1 3215  
 1 MluI Acgct 1 3527  
 1 HpaI GTTaa 1 3730  
 1 XbaI Tctaga 1 3767  
 1  
 1 AfII Cttaag 1 3811  
 1 BsmI NGcattc 1 3821  
 1 "- GAATGCN 1 4695  
 1 RsrII CGgwccg 1 3827  
 1 NheI Gctagc 1 4166  
 1 BstEII Ggtnac 1 4182  
 1 BsmBI CGTCTCNnnnn 2 4188 6625  
 1 "- Nnnnnngagacg 1 6673  
 1 ApaI GGGCCc 1 4209  
 1 BanII GRGCYc 3 4209 4492 6319  
 1 BspI20I Gggccc 1 4209  
 1 PspOMI Gggccc 1 4209  
 1 BseRI NNNnnnnnnctctc 1 4226  
 1 "- GAGGAGNNNNNNNNNN 1 4957  
 1 EcoNI CCTNNnnnagg 1 4278  
 1 PfuI GACNngtc 1 4308  
 1 TthI11I GACNngtc 1 4308





```

5      61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
      E E R F P M M S T F K V L L C
      381 gaa gaa cgt ttt cca atg atg agc act ttt aaa gtt ctg cta tgt

      76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
      G A V L S R I D A G Q E Q L G
      426 ggc ggc gta tta tcc cgt att gac gcc ggc caa gaG CAa ctc ggT
      Bgl.....

10     91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
      R R I H Y S Q N D L V E Y S P
      471 CGc cgc ata cac tat tct cag aat gac tgg gtt gAG TAC Tca cca
      Bgl.....
      Scal...

15     106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
      V T E K H L T D G M T V R E L
      516 gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga gaa tta

      121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
      C S A A I T M S D N T A A N L
      561 tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac tta

20     136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
      L L T T I G G P K E L T A F L
      606 ctt ctg aca aCGATC Gga gga ccg aag gag cta acc gct ttt tgg
      PvuI.... (1/2)

25     151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
      H N M G D H V T R L D R W E P
      651 cac aac atg ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg

30     166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
      E L N E A I P N D E R D T T M
      696 gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg

```

[illegible]

```

1081 cctatatact tagatgat ttaaaactt attttaatt taaaaggatc taggtgaaga
1141 tcttttga taattcatg accaaatcc cttaacgtga gttttcgtc cactgaaggt
1201 cagaccocgt agaaagcgt aaaggatctt ctgagatcc ttttttcg cgcgtaact
1261 gcgtcttga aacaaaanaa ccaccgtac cagcggfagt ttgttgcg gatcangagc
1321 taccactct ttttcegaag gtaactgggt tcagcagagc gcagatacca aatactgtcc
1381 ttctagtta gccgtagtta ggccaccact tcaagaactc tglugcaocg cctacatacc
1441 tcgtcttgt aatctctgta ccagtggcgt ctgccagtgg cgataagtcg tgcitaceg
1501 ggttggacit aagcagatag ttaccggata aggcgcagcg gtcgggctga acggggggtt
1561 cglcatala gccagcttg gagcaacga cctacacga actgagatac ctacagcgtg
1621 agcatfaga aagcgcacg ctcccgag aggaaaggc ggacagGTAT OCggttaagcg
      BclVI. (2 of 2)
1681 gcagagtcgg aacagagag cgCAGCAGgg agcttcagg gggaaagcc tgglatctt
      BssSI(2/2)
1741 atagctctgt cgggtttcgc cactctgac ttgagctcg atttttga tgcctgtcag
1801 gggggcggag cctatggaa aagccagca acgcgccct ttacgggtc ctggccittt
1861 gctggccctt tctcACATG Ttcttctcg cgttatccc tgaattcgt gataaccgta
      PciI...
1921 ttaccgctt tgaagtgtat gataccgtc gccgcagcg aacgaccgag cgcagcgagt
1981 cagtgagcga ggaagcgGAA GAGCgoccaa tacgcaaac gcctctccc gcgcgttggc
      SapI...
2041 cgattcatia atgCAGCTGg cagcagaggt ttcccgactg gaaagcgggc agtgagcgca
      PvuII(1/3)
2101 acgcaatTAA TGTgagtiag ctactcatt aggcacccca ggcTTTACac ttatgcttc
      ..-35.. Plac
2161 cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacaCAGGA AACAGCTATG
      M13Rev_seq_primer
2221 ACcatgata cgCCAAAGCTT TGGagccctt ttttggaga ttttcaac
      PflMI.....
      Hind3.
! signal::linker::CLight
!
! 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
! fM K L L F A I P L V P F Y
2269 gtg aaa aaa tia tta ttc gca att cct tta gtt cct cct ttc tat
!
35 ! Linker ..... End of FR4

```

```

16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
S H S A Q V Q L Q V D L E I K
2314 tct cac aGT GCA Cag gtc caa CTG CAG GTC GAC CTC GAG atc aaa
      ApaLI..... PstI... XhoI...
5      BspMI...
      Sall...
      AccI...(1/2)
      HincII.(1/2)

10  Vlight domains could be cloned in as ApaLI-XhoI fragments.
     VL-CL(kappa) segments can be cloned in as ApaLI-AscI fragments <-----

31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
R G T V A A P S V F I F P P S
2359 cgt gga act gtg gct gca cca tct GTC TTC atc ttc ccg cca tct
      BbsI...(1/2)

46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
D E Q L K S G T A S V V C L L
2404 gat gag cag ttg aaa tct gga act gcc tct gtt gtg tgc ctg ctg

61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
N N F Y P R E A K V Q W K V D
2449 aat aac ttc tat ccc aga gag gcc aaa gta cag tgg aag gtg gat

76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
N A L Q S G N S Q E S V T E Q
2494 aac gcc ctc caa tog ggt aac tcc cag gag agt gtc aca gag cag

91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
D S K D S T Y S L S T L T L
2539 gac agc gag acc tac agc ctc agc agc acc ctg acG CTG
      EspI...

106 107 108 109 110 111 112 113 114 115 116 117 118 119 120

```

- 228 -

```

1   S K A D Y E K H K V Y A C E V
2584 AGC aag gca tac gag aac cag aac GTC TAC gcc tgc gaa gtc
    ...Epl...
    Acl... (2/2)
5   121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
    T H Q G L S S P V T K S F N R
2629 acc cat cag ggc ctg agt tca CCG GTg aca aag agc ttc aac agg
    AgeI... (1/2)
10  136 137 138 139 140
    G E C .
2674 gga gag tgt taa taa GG CGCGCCaatt
    AscI....
    BssHII.
15  2701 ctattcaag gagacagta ta
    PelB::3-23(stuffed)::CH1::III fusion gene
20  1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
    M K Y L L P T A A A G L L L L
2723 atg aag tac cta ttg cct acg gca gcc gct gga ttg tta tta ctc
    -----
25  16 17 18 19 20 21 22
    A A Q P A M A
2768 gcG GCC cag ccG GCC atg gcc
    SfiI.....
    NgoMIV.. (1/2)
    NcoI...
30  FR1(DP47/V3-23)-----
    23 24 25 26 27 28 29 30
    E V Q L L E S G
    gaa|gtt|CAA|TTG|tta|gag|tct|ggg|
35  2789

```

5	31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 G G L V Q P G S L R L S C A 2813  ggc ggctt gtt cag cct ggc t ggc t c t a t c g t c t t c g c t	MfeI
10	----FR1----- 46 47 48 A S G 2858  gct TCC GGA   BspEI	
15	-----FR2-----> There are no stop codons in this staffer. 2867  gctt c GATC gctt	
20	BelII. 2887  tut gggg ggc cagatc ggc ttacgga gatc gaaccga ctgct ggac aaagaccag  2947  ctta tgc T GATC Agcatt ggg atgttat tcgcc aaacc agt gtc cagg atct taacct	
25	XcmI..... 3247  cga aacacc cagg acggcc caact ggctc gct gaalata agtg ttggag caaaaattt  3307  gat gaggcg gtgc caggcg acaa taccac antcccacag gcgg ttgatc t gtt ct c g  3367  gaa accacag cagg agggtg tg tgg ctgc gct gggaagat acc t ggagaga ctct ttccaa  3427  acgt atggc aat aat tga gta act tggaa aac act tga atgg cct t aa cg tcc ggcg  3487  aa taattc ttg tgt ac cgc agg ccgc agc cga agaa ACGGCT c ac agg c ggag ta	
30	Mul. 3547  taaa accgt ggac acagaa acg ata gat t gtt cttca cca acg acaa gcg atc gtcc  3607  tg ct tccc tggg at tgg tcg acc ccgc tcag at tggg ttat t gct ccg at gg aac  3667  agt g taag cac atg aag atc agt tga aat g tac gaa aat tt ggcc g ta at t c gt	
35	PvuII. 3727  ctg GTTAAC g aag cag gatg tgg agg cgca taagg atc g	





- 231 -

```

1 | K S T S G G T A A L G C L V K
4243 | aag agc acc tct ggg ggc aca ggc gcc ctg ggc ctg gtc aag
5 | 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
   | D Y F P E P V T V S W N S G A
4288 | gac tac ttc ccc gaa cgg gfg acg gfg tgg aac tca ggc gcc
10 | 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
   | L T S G V H T F P A V L Q S S
4333 | ctg acc agc ggc gtc cac acc ttc cgg gct gtc cta cag tcc tca
   | 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
   | G L Y S L S V V T V P S S S
4378 | gga ctc tac tcc ctc agc agc gta gfg acc gfg ccc tcc agc agc
15 | 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225
   | L G T Q T Y I C N V N H K P S
4423 | ttg ggc acc cag acc tac atc tgc aac gfg aat cac aag ccc agc
   | 226 227 228 229 230 231 232 233 234 235 236 237 238
   | N T K V D K K V E P K S C
20 | 4468 | aac ac aag gfg gac aaG AAA GTT GAG CCC AAA TCT TGT
   | ON-TQHCforw.....
   | Poly His linker
25 | 139 140 141 142 143 144 145 146 147 148 149 150
   | A A A H H H H H G A A
   | GCG GCC GCA cat cat cat cac cat cac ggg gcc gca
4507 | Notl.....
   | Eagl....
30 | 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
   | E Q K L I S E D L N G A A
4543 | gaa caa aaa ctc atc tca gaa gag gat ctg aat ggg gcc gca tag
35 | Mature III----->...

```

- 232 -

```

166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
T V E S C L A K P H T E N S F
4588 act gti gaa agt tta gca aaa cct cat aca gaa aat tca tti
!
181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
T N V W K D D K T L D R Y A N
4633 act aac gtc tgg aaa gac aac act tta gat cgt tac gct aac
!
196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
Y E G C L W N A T G V V V C T
4678 tat gag ggc tgt ctg tgG AAT GCt aca ggc gti gfg gtt tgt act
!
BsmI....
211 212 213 214 215 216 217 218 219 220 221 222 223 224 225
G D E T Q C Y G T W V P I G L
4723 ggt gac gaa act cag tct tac ggt aca tgg gtt cct att ggc ctt
!
226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
A I P E N E G G S E G G S
4768 gct atc cct gaa aat gag ggt ggt ggc tct gag ggt ggc ggt tct
!
241 242 243 244 245 246 247 248 249 250 251 252 253 254 255
E G G S E G G T K P P E Y
4813 gag ggt ggc ggt tct gag ggt ggc ggt act aaa cct cct gag tac
!
256 257 258 259 260 261 262 263 264 265 266 267 268 269 270
G D T P I P G Y T Y I N P L D
4858 ggt gat aca cct att ceg ggc tat act tat atc aac cct ctc gac
!
271 272 273 274 275 276 277 278 279 280 281 282 283 284 285
G T Y P P G T E Q N P A N P N
4903 ggc act tat ceg cct ggt act gag caa aac ccc gct aat cct aat
!

```

```

!      286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
!      P S L E S Q P L N T F M F Q
! 4948 cct ctt GAG GAG tct cag cct ctt aat act ttc atg ttt cag
!      BseRI.(2/2)
!
!      301 302 303 304 305 306 307 308 309 310 311 312 313 314 315
!      N N R F R N R Q G A L T V Y T
! 4993 aat aat agg ttc cga aat agg cag ggt gca tia act gtt tat acg
!
!      316 317 318 319 320 321 322 323 324 325 326 327 328 329 330
!      G T V T Q G T D P V K T Y Y Q
! 5038 ggc act gtt act caa ggc act gac ccc gtt aaa act tat tac cag
!
!      331 332 333 334 335 336 337 338 339 340 341 342 343 344 345
!      Y T P V S S K A M Y D A Y W N
! 5083 tac act cct gta tca tca aaa gcc alg tat gac gct tac tgg aac
!
!      346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
!      G K F R D C A F H S G F N E D
! 5128 ggt aaa ttc aga gac tgc gct ttc cat tct ggc ttt aat gaG GAT
!      BamHI..
!
!      361 362 363 364 365 366 367 368 369 370 371 372 373 374 375
!      P F V C E Y Q G Q S S D L P Q
! 5173 CCa ttc gtt tgt gaa tat caa ggc caa tgc tct gAC CTG Cct caa
!      BamHI..
!      BspMI..(2/2)
!
!      376 377 378 379 380 381 382 383 384 385 386 387 388 389 390
!      P P V N A G G S G G S G G
! 5218 cct cct gtc aat gct ggc ggc tct ggt ggt ggt tct ggt ggc
!
!      391 392 393 394 395 396 397 398 399 400 401 402 403 404 405
!      G S E G G S E G G S E G G
! 5263 ggc tct gag ggt ggc ggc tct gag ggt ggc ggt tct gag ggt ggc
!
!      406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
!

```

```

1  G S E G G S G G S G S G D
5308 ggc tct gag ggt ggc ggc tcc ggt ggc ggc tcc ggt tcc ggt gat
1
2  421 422 423 424 425 426 427 428 429 430 431 432 433 434 435
   F D Y E K M A N A N K G A M T
5  5353 ttg gat tat gaa aaa atg gca aac gct aat aag ggg gct atg acc
1
6  436 437 438 439 440 441 442 443 444 445 446 447 448 449 450
   E N A D E N A L Q S D A K G K
10 5398 gaa aat gcc gat gaa aac ggc cta cag tct gac gct aaa ggc aaa
1
11 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465
   L D S V A T D Y G A A I D G F
15 5443 ctt gat tct gtc gct act gat tac ggt gct gct ATC GAT ggt ttc
   BspDI..
1
16 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
   I G D V S G L A N G N G A T G
20 5488 att ggt gac ggt tcc ggc ctt gct aat ggt aat ggt gct act ggt
1
21 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495
   D F A G S N S Q M A Q V G D G
25 5533 gat ttg gct ggc tct aat tcc caa atg gct caa gtc ggt gac ggt
1
26 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510
   D N S P L M N N F R Q Y L P S
30 5578 gat aat tca cct tta atg aat aat ttc cgt caa tat tia cct tct
1
31 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525
   L P Q S V E C R P Y V F G A G
35 5623 ttg cct cag tgc ggt gaa tgt ogc cct tat gtc ttt ggc gct ggt
1
36 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540
   K P Y E F S I D C D K I N L F
37 5668 aaa cCA TAT Gaa ttg att ggt gac aaa ala aac tia ttc
   NdeI..

```

```

!      541 542 543 544 545 546 547 548 549 550 551 552 553 554 555
!      R G V F A F L L Y V A T F M Y
!      5713 cgt ggc ttt ggc ttt ctt tta tat gtt gcc acc ttt atg tat
!
!      556 557 558 559 560 561 562 563 564 565 566 567 568 569 570
!      V F S T F A N I L R N K E S
!      5758 gta ttt tgc acg ttt gct aac ata cta cta cgt aat aag gag tct taa
!
!      571
!
!      5803 taa GAATTC
!      EcoRI
!      5812 actggccgt cgttttacaa cgtcgtgact gggaataacc tggcgttacc caactaatc
!      5871 gccctgcgc acatccccc ttgcacagct ggctgaatag cgaagaggcc cgcacCGATC
!      PvuI..
!      5931 Gcccttccca acagtTGGCG AGcctgaatg gegaatGGCG CCtgaigcgg tatttctcc
!      ...PvuI... (3/3) FspI.. (2/2) KasI... (2/2)
!      5991 ttacgcact gtgcggta ttacaccgca tataaatgt aaacgttaat atttggtaa
!      6051 aaatcggtt aaattttgt taatcagct caitttttaa ccaatagcc gaaatcgcca
!      6111 aaatcccTTA TAAatcaana gaatagccg agatagggtt gagtgtgtt ccagtttggga
!      PstI...
!      6171 acaagagtc actattaaag aacgtggact ccaacgtcaa agggcgaaaa acgctctac
!      6231 agggcgatgg ccCACtacGT Gaaccatcac ccaatcaag tttttgggg tggagggtcc
!      DraIII....
!      6291 gtaagcact aaatcggaac cctaaaggga gcccccgatt tagagcttga cggggaaaGC
!      NgoMIV..
!      6351 CGGCGaacgt ggcgagaaag gaagggaaga aagcgaaagg agcgggcgt agggcgctgg
!      ..NgoMIV.(2/2)
!      6411 caagtgtac ggtcacgctg cgcglaaaca ccacaccgc cgcgcttaat gcgcgcctac
!      6471 agggcggtga ctatgttgc ttggacgggt gcagctctag tacaatctgc tctatgcgc
!      6531 catagttaag ccagcccgga caccgcccaa caccgcctga cgcgccttgc cgggcttgc
!      6591 tgcctccgc atcgcttac agacaagctg tgaccgtct cgggagctgc atgtgtcaga
!      6651 gggtttcaac gtcataccg aaacgcgga

```

Table 30: Oligonucleotides used to clone CDR1/2 diversity

All sequences are 5' to 3'.

1) ON\_CD1Bsp, 30 bases

5 A c c T c A c T g g c T T c c g g A  
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18  
T T c A c T T T c T c T  
10 19 20 21 22 23 24 25 26 27 28 29 30

2) ON\_Br12, 42 bases

15 A g A A A c c c A c T c c A A A c c  
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18  
T T T A c c A g g A g c T T g g c g  
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36  
20 A A c c c A  
37 38 39 40 41 42

3) ON\_CD2Xba, 51 bases

25 g g A A g g c A g T g A T c T A g A  
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18  
g A T A g T g A A g c g A c c T T T  
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36  
30 A A c g g A g T c A g c A T A  
37 38 39 40 41 42 43 44 45 46 47 48 49 50 51

35 4) ON\_BotXba, 23 bases

g g A A g g c A g T g A T c T A g A  
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18  
40 g A T A g  
19 20 21 22 23

Table 31: Bridge/Extender Oligonucleotides

	ON_Lam1aB7(rc)	.....GTGCTGACTCAGCCACCCTC.	20
	ON_Lam2aB7(rc)	.....GCCCTGACTCAGCCTGCCTC.	20
5	ON_Lam31B7(rc)	.....GAGCTGACTCAGG.ACCCTGC	20
	ON_Lam3rB7(rc)	.....GAGCTGACTCAGCCACCCTC.	20
	ON_LamHf1cBrg(rc)	CCTCGACAGCGAAGTGCACAGAGCGTCTTGACTCAGCC.....	38
	ON_LamHf1cExt	CCTCGACAGCGAAGTGCACAGAGCGTCTTG.....	30
	ON_LamHf2b2Brg(rc)	CCTCGACAGCGAAGTGCACAGAGCGCTTTGACTCAGCC.....	38
10	ON_LamHf2b2Ext	CCTCGACAGCGAAGTGCACAGAGCGCTTTG.....	30
	ON_LamHf2dBrg(rc)	CCTCGACAGCTAAGTGCACAGAGCGCTTTGACTCAGCC.....	38
	ON_LamHf2dExt	CCTCGACAGCGAAGTGCACAGAGCGCTTTG.....	30
	ON_LamHf31Brg(rc)	CCTCGACAGCGAAGTGCACAGAGCGAATTGACTCAGCC.....	38
	ON_LamHf31Ext	CCTCGACAGCGAAGTGCACAGAGCGAATTG.....	30
15	ON_LamHf3rBrg(rc)	CCTCGACAGCGAAGTGCACAGTACGAATTGACTCAGCC.....	38
	ON_LamHf3rExt	CCTCGACAGCGAAGTGCACAGTACGAATTG.....	30
	ON_lamPlePCR	CCTCGACAGCGAAGTGCACAG.....	21
	Consensus		

Table 32: Oligonucleotides used to make SSDNA locally  
double-stranded

Adapters (8)	
5	H43HF3.1-02#1 5'-cc gtg tat tac tgt gcg aga g-3'
	H43.77.97.1-03#2 5'-ct gtg tat tac tgt gcg aga g-3'
	H43.77.97.323#22 5'-cc gta tat tac tgt gcg aaa g-3'
	H43.77.97.330#23 5'-ct gtg tat tac tgt gcg aaa g-3'
	H43.77.97.439#44 5'-ct gtg tat tac tgt gcg aga c-3'
10	H43.77.97.551#48 5'-cc atg tat tac tgt gcg aga c-3'



Table 33: Bridge/extender pairs

Bridges (2)

H43.XABr1

- 5 5'ggtgtagtgaTCTAGtgacaactctaagaatactctctacttgagatgaacagC  
TTtAGggctgaggacaCTGCAGtctactattgtgcgaga-3'

H43.XABr2

- 5'ggtgtagtgaTCTAGtgacaactctaagaatactctctacttgagatgaacagC  
10 TTtAGggctgaggacaCTGCAGtctactattgtgcgaaa-3'

Extender

H43.XAExt

- 5'ATAgTAgAcTgcAgTgTccTcAgcccTTAAgcTgTTcATcTgcAAgTAgAgAgTA  
15 TTcTTAgAgTTgTcTcTAgATcAcTAcAcc-3'

10045674-100504

Table 34: PCR primers

Primers

H43.XAPCR2	gactgggTgTAgTgATcTAg
5 Hucmnest	cttttctttgttgccggtggggtg

Table 35: PCR program for amplification of heavy chain CDR3 DNA

5	95 degrees C	5 minutes	
	95 degrees C	20 seconds	
	60 degrees C	30 seconds	repeat 20x
	72 degrees C	1 minute	
10	72 degrees C	7 minutes	
	4 degrees C	hold	

Reagents (100 ul reaction):

	Template	5ul ligation mix
15	10x PCR buffer	1x
	Taq	5U
	dNTPs	200 uM each
	MgCl <sub>2</sub>	2mM
	H43.XAPCR2-biotin	400 nM
20	Hucmnest	200 nM

! Table 36: Annotated sequence of CJR DY3F7(CJR-A05) 10251 bases

! Non-cutters					
5	!BclI Tgatca	!BsiWI Cgtacg	!BssSI Cacgag		
	!BstZ17I GTAtac	!BtrI CACgtg	!EcoRV GATatc		
	!FseI GGCCGGcc	!HpaI GTTaac	!MluI Acgcgt		
	!PmeI GTTTaaac	!PmlI CACgtg	!PpuMI RGgwccy		
	!RsrII CGgwccg	!SapI GCTCTTC	!SexAI Accwggg		
10	!SgfI GCGATcgc	!SgrAI CRccggyg	!SphI GCATGc		
	!StuI AGGcct	!XmaI Cccggg			
! cutters					
15	! Enzymes that cut from 1 to 4 times and other features				
	!End of genes II and X	829			
	!Start gene V	843			
	!BsrGI Tgtaca	1	1021		
20	!BspMI Nnnnnnnnngcaggt	3	1104	5997	9183
	!-"- ACCTGCNNNNn	1	2281		
	!End of gene V		1106		
	!Start gene VII		1108		
	!BsaBI GATNNnnatc	2	1149	3967	
25	!Start gene IX		1208		
	!End gene VII		1211		
	!SnaBI TACgta	2	1268	7133	
	!BspHI Tcatga	3	1299	6085	7093
	!Start gene VIII		1301		
30	!End gene IX		1304		
	!End gene VIII		1522		
	!Start gene III		1578		
	!EagI Cggccg	2	1630	8905	
	!XbaI Tctaga	2	1643	8436	
35	!KasI Ggcgcc	4	1650	8724	9039 9120
	!BsmI GAATGCN	2	1769	9065	
	!BseRI GAGGAGNNNNNNNNNN	2	2031	8516	
	!-"- NNnnnnnnnnctcctc	2	7603	8623	
	!AlwNI CAGNNNctg	3	2210	8072	8182
40	!BspDI ATcgat	2	2520	9883	
	!NdeI CATatg	3	2716	3796	9847
	!End gene III		2846		
	!Start gene VI		2848		
	!AfeI AGCgct	1	3032		
45	!End gene VI		3187		
	!Start gene I		3189		
	!EarI CTCTTCNnnn	2	4067	9274	
	!-"- Nnnngaagag	2	6126	8953	
	!PacI TTAATtaa	1	4125		
50	!Start gene IV		4213		
	!End gene I		4235		
	!BsmFI Nnnnnnnnnnnnnngtccc	2	5068	9515	
	!MscI TGGcca	3	5073	7597	9160
	!PsiI TTAtaa	2	5349	5837	
55	!End gene IV		5493		
	!Start ori		5494		
	!NgoMIV Gccggc	3	5606	8213	9315
	!BanII GRGCYc	4	5636	8080	8606 8889
	!DraIII CACNNNgtg	1	5709		
60	!DrdI GACNNNNngtc	1	5752		
	!AvaI Cycgrg	2	5818	7240	
	!PvuII CAGctg	1	5953		

	!BsmBI CGTCTCnNNnn	3	5964	8585	9271
	!End ori region		5993		
	!BamHI Ggatcc	1	5994		
	!HindIII Aagctt	3	6000	7147	7384
5	!BciVI GTATCCNNNNNN	1	6077		
	!Start bla		6138		
	!Eco57I CTGAAG	2	6238	7716	
	!SpeI Actagt	1	6257		
	!BcgI gcannnnnnntcg	1	6398		
10	!ScaI AGTact	1	6442		
	!PvuI CGATcg	1	6553		
	!FspI TGCgca	1	6700		
	!BglI GCCNNNNNnggc	3	6801	8208	8976
	!BsaI GGTCTCnNNnn	1	6853		
15	!AhdI GACNNNNngtc	1	6920		
	!Eam1105I GACNNNNngtc	1	6920		
	!End bla		6998		
	!AccI GTmkac	2	7153	8048	
	!HincII GTYrac	1	7153		
20	!SalI Gtcgac	1	7153		
	!XhoI Ctcgag	1	7240		
	!Start PlacZ region		7246		
	!End PlacZ region		7381		
	!PflMI CCANNNNNntgg	1	7382		
25	!RBS1		7405		
	!start M13-iii signal seq for LC		7418		
	!ApaLI Gtgcac	1	7470		
	!end M13-iii signal seq		7471		
	!Start light chain kappa L20:JK1		7472		
30	!PflFI GACNnnngtc	3	7489	8705	9099
	!SbfI CCTGCAGg	1	7542		
	!PstI CTGCAG	1	7543		
	!KpnI GGTACc	1	7581		
	!XcmI CCANNNNNnnnnntgg	2	7585	9215	
35	!NsiI ATGCAc	2	7626	9503	
	!BsgI ctgcac	1	7809		
	!BbsI gtcttc	2	7820	8616	
	!BlpI GCTnagc	1	8017		
	!EspI GCTnagc	1	8017		
40	!EcoO109I RGgnccy	2	8073	8605	
	!Ecl136I GAGctc	1	8080		
	!SacI GAGCTc	1	8080		
	!End light chain		8122		
	!AscI GGcgcgcc	1	8126		
45	!BssHII GcgcgC	1	8127		
	!RBS2		8147		
	!SfiI GGCCNNNNNnggcc	1	8207		
	!NcoI Ccatgg	1	8218		
	!Start 3-23, FR1		8226		
50	!MfeI Caattg	1	8232		
	!BspEI Tccgga	1	8298		
	!Start CDR1		8316		
	!Statt FR2		8331		
	!BstXI CCANNNNNntgg	2	8339	8812	
55	!EcoNI CCTNNnnnagg	2	8346	8675	
	!Start FR3		8373		
	!XbaI Tctaga	2	8436	1643	
	!AflIII Cttaag	1	8480		
	!Start CDR3		8520		
60	!AatII GACGTc	1	8556		
	!Start FR4		8562		
	!PshAI GACNNnnngtc	2	8573	9231	

```

!BstEII Ggtnacc          1      8579
!Start CH1                8595
!ApaI GGGCCc             1      8606
!Bsp120I Gggccc          1      8606
5  !PspOMI Gggccc          1      8606
!AgeI Accggt             1      8699
!Bsu36I CCtnagg          2      8770    9509
!End of CH1              8903
!NotI GCggccgc           1      8904
10 !Start His6 tag         8913
!Start cMyc tag           8931
!Amber codon              8982
!NheI Gctagc             1      8985
!Start M13 III Domain 3   8997
15 !NruI TCGcga           1      9106
!BstBI TTcgaa            1      9197
!EcoRI Gaattc            1      9200
!XcmI CCANNNNNnnntgg     1      9215
!BstAPI GCANNNNNntgc     1      9337
20 !SacII CCGCgg          1      9365
!End IIIstump anchor     9455
!AvrII Cctagg            1      9462
!trp terminator          9470
!SwaI ATTTaaat          1      9784
25 !Start gene II         9850
!BglIII Agatct           1      9936
!-----
--
1 aat gct act act att agt aga att gat gcc acc ttt tca gct cgc
30 gcc
! gene ii continued
49 cca aat gaa aat ata gct aaa cag gtt att gac cat ttg cga aat
gta
97 tct aat ggt caa act aaa tct act cgt tcg cag aat tgg gaa tca
35 act
145 gtt aTa tgg aat gaa act tcc aga cac cgt act tta gtt gca tat
tta
193 aaa cat gtt gag cta cag caT TaT att cag caa tta agc tct aag
cca
40 241 tcc gca aaa atg acc tct tat caa aag gag caa tta aag gta ctg
tct
289 aat cct gac ctg ttg gag ttt gct tcc ggt ctg gtt cgc ttt gaa
gct
337 cga att aaa acg cga tat ttg aag tct ttc ggg ctt cct ctt aat
45 ctt
385 ttt gat gca atc cgc ttt gct tct gac tat aat agt cag ggt aaa
gac
433 ctg att ttt gat tta tgg tca ttc tcg ttt tct gaa ctg ttt aaa
gca
50 481 ttt gag ggg gat tca ATG aat att tat gac gat tcc gca gta ttg
gac
! Start gene x, ii continues
529 gct atc cag tct aaa cat ttt act att acc ccc tct ggc aaa act
tct
55 577 ttt gca aaa gcc tct cgc tat ttt ggt ttt tat cgt cgt ctg gta
aac
625 gag ggt tat gat agt gtt gct ctt act atg cct cgt aat tcc ttt
tggt
673 cgt tat gta tct gca tta gtt gaa tgt ggt att cct aaa tct caa
60 ctg
721 atg aat ctt tct acc tgt aat aat gtt gtt ccg tta gtt cgt ttt
att

```

```

769 aac gta gat ttt tct tcc caa cgt cct gac tgg tat aat gag cca
gtt
817 ctt aaa atc gca TAA
!                               End X & II
5 832 ggtaattca ca
!
!           M1           E5           Q10           T15
843 ATG att aaa gtt gaa att aaa cca tct caa gcc caa ttt act act
cgt
10 !       Start gene V
!
!           S17           S20           P25           E30
891 tct ggt gtt tct cgt cag ggc aag cct tat tca ctg aat gag cag
ctt
15 !
!           V35           E40           V45
939 tgt tac gtt gat ttg ggt aat gaa tat ccg gtt ctt gtc aag att
act
!
20 !           D50           A55           L60
987 ctt gat gaa ggt cag cca gcc tat gcg cct ggt cTG TAC Acc gtt
cat
!
!           L65           V70           S75
25 R80
1035 ctg tcc tct ttc aaa gtt ggt cag ttc ggt tcc ctt atg att gac
cgt
!
!           P85           K87 end of V
30 1083 ctg cgc ctc gtt ccg gct aag TAA C
!
!       1108 ATG gag cag gtc gcg gat ttc gac aca att tat cag gcg atg
!       Start gene VII
!
35 1150 ata caa atc tcc gtt gta ctt tgt ttc gcg ctt ggt ata atc
!
!           VII and IX overlap.
!           ..... S2 V3 L4 V5           S10
1192 gct ggg ggt caa agA TGA gt gtt tta gtg tat tct ttT gcc tct ttc
40 gtt
!
!           End VII
!           |start IX
!           L13           W15           G20           T25
E29
45 1242 tta ggt tgg tgc ctt cgt agt ggc att acg tat ttt acc cgt tta
atg gaa
!
1293 act tcc tc
!
50 !       .... stop of IX, IX and VIII overlap by four bases
1301 ATG aaa aag tct tta gtc ctc aaa gcc tct gta gcc gtt gct acc
ctc
!       Start signal sequence of viii.
!
55 1349 gtt ccg atg ctg tct ttc gct gct gag ggt gac gat ccc gca aaa
gcg
!
!           mature VIII --->
1397 gcc ttt aac tcc ctg caa gcc tca gcg acc gaa tat atc ggt tat
gcg
60 1445 tgg gcg atg gtt gtt gtc att
1466 gtc ggc gca act atc ggt atc aag ctg ttt aag
!
```

```

! bases 1499-1539 are probable promoter for iii
1499 aaa ttc acc tcg aaa gca ! 1515
! ..... -35 ..
5 1517 agc tga taaaccgat acaattaaag gctccttttg
! ..... -10 ...
1552 gagccttttt ttt GGAGAt ttt ! S.D. uppercase, there may be 9 Ts
10 ! <----- III signal sequence ----->
! M K K L L F A I P L V V P F
1574 caac GTG aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc !
1620
15 ! Y S G A A E S H L D G A
1620 tat tct ggc gCG GCC Gaa tca caT CTA GAc ggc gcc
! EagI.... XbaI....
20 ! Domain 1 -----
! A E T V E S C L A
1656 gct gaa act gtt gaa agt tgt tta gca
25 ! K S H T E I S F T N V W K D D K
T
1683 aaA Tcc cat aca gaa aat tca ttc aCT AAC GTC TGG AAA GAC GAC
AAA ACT
30 ! L D R Y A N Y E G S L W N A T G
V
1734 tta gat cgt tac gct aac tat gag ggC tgt ctg tgG AAT GCt aca
ggc gtt
! BsmI....
35 ! V V C T G D E T Q C Y G T W V P
I
1785 gta gtt tgt act ggt GAC GAA ACT CAG TGT TAC GGT ACA TGG GTT
cct att
40 ! G L A I P E N
1836 ggg ctt gct atc cct gaa aat
!
! L1 linker -----
45 ! E G G G S E G G G S
1857 gag ggt ggt ggc tct gag ggt ggc ggt tct
!
! E G G G S E G G G T
1887 gag ggt ggc ggt tct gag ggt ggc ggt act
50 !
! Domain 2 -----
1917 aaa cct cct gag tac ggt gat aca cct att ccg ggc tat act tat
atc aac
1968 cct ctc gac ggc act tat ccg cct ggt act gag caa aac ccc gct
aat cct
55 2019 aat cct tct ctt GAG GAG tct cag cct ctt aat act ttc atg ttt
cag aat
! BseRI..
2070 aat agg ttc cga aat agg cag ggg gca tta act gtt tat acg ggc
act
60 2118 gtt act caa ggc act gac ccc gtt aaa act tat tac cag tac act
cct

```

T004694.10604



```

2166 gta tca tca aaa gcc atg tat gac gct tac tgg aac ggt aaa ttC
AGA
!
AlwNI
5 2214 GAC TGc gct ttc cat tct ggc ttt aat gaG gat TTa ttT gtt tgt
gaa
! AlwNI
2262 tat caa ggc caa tcg tct gac ctg cct caa cct cct gtc aat gct
!
10 2307 ggc ggc ggc tct
! start L2 -----
-----
2319 ggt ggt ggt tct
2331 ggt ggc ggc tct
15 2343 gag ggt ggt ggc tct gag gga ggc ggt tcc
2373 ggt ggt ggc tct ggt ! end L2
!
! Many published sequences of M13-derived phage have a longer linker
! than shown here by repeats of the EGGGS motif two more times.
20 !
! Domain 3 -----
-----
! S G D F D Y E K M A N A N K G A
2388 tcc ggt gat ttt gat tat gaa aag atg gca aac gct aat aag ggg
25 gct
!
! M T E N A D E N A L Q S D A K G
2436 atg acc gaa aat gcc gat gaa aac gcg cta cag tct gac gct aaa
ggc
30 !
! K L D S V A T D Y G A A M D G F
2484 aaa ctt gat tct gtc gct act gat tac ggt gct gct atc gat ggt
ttc
!
! I G D V S G L A N G N G A T G D
35 2532 att ggt gac gtt tcc ggc ctt gct aat ggt aat ggt gct act ggt
gat
!
! F A G S N S Q M A Q V G D G D N
40 2580 ttt gct ggc tct aat tcc caa atg gct caa gtc ggt gac ggt gat
aat
!
! S P L M N N F R Q Y L P S L P Q
2628 tca cct tta atg aat aat ttc cgt caa tat tta cct tcc ctc cct
45 caa
!
! S V E C R P F V F G A G K P Y E
2676 tcg gtt gaa tgt cgc cct ttt gtc ttt Ggc gct ggt aaa cca tat
gaa
50 !
! F S I D C D K I N L F R
2724 ttt tct att gat tgt gac aaa ata aac tta ttc cgt
! End Domain 3
!
! G V F A F L L Y V A T F M Y V
55 F140
2760 ggt gtc ttt gcg ttt ctt tta tat gtt gcc acc ttt atg tat gta
ttt
! start transmembrane segment
60 !
! S T F A N I L
2808 tct acg ttt gct aac ata ctg

```

```

!
!   R   N   K   E   S
2829 cgt aat aag gag tct TAA ! stop of iii
!   Intracellular anchor.
5
!   M1  P2  V   L  L5   G   I   P   L  L10  L   R   F   L
G15
2847 tc ATG cca gtt ctt ttg ggt att ccg tta tta ttg cgt ttc ctc
ggt
10
!   Start VI
!
2894 ttc ctt ctg gta act ttg ttc ggc tat ctg ctt act ttt ctt aaa
aag
2942 ggc ttc ggt aag ata gct att gct att tca ttg ttt ctt gct ctt
15
att
2990 att ggg ctt aac tca att ctt gtg ggt tat ctc tct gat att agc
gct
3038 caa tta ccc tct gac ttt gtt cag ggt gtt cag tta att ctc ccg
tct
20
3086 aat gcg ctt ccc tgt ttt tat gtt att ctc tct gta aag gct gct
att
3134 ttc att ttt gac gtt aaa caa aaa atc gtt tct tat ttg gat tgg
gat
25
!   M1  A2  V3           F5                               L10           G13
3182 aaa TAA t ATG gct gtt tat ttt gta act ggc aaa tta ggc tct gga
!   end VI   Start gene I
!
!   K   T   L   V   S   V   G   K   I   Q   D   K   I   V   A
30
3228 aag acg ctc gtt agc gtt ggt aag att cag gat aaa att gta gct
!
!   G   C   K   I   A   T   N   L   D   L   R   L   Q   N   L
3273 ggg tgc aaa ata gca act aat ctt gat tta agg ctt caa aac ctc
!
35
!   P   Q   V   G   R   F   A   K   T   P   R   V   L   R   I
3318 ccg caa gtc ggg agg ttc gct aaa acg cct cgc gtt ctt aga ata
!
!   P   D   K   P   S   I   S   D   L   L   A   I   G   R   G
3363 ccg gat aag cct tct ata tct gat ttg ctt gct att ggg cgc ggt
40
!
!   N   D   S   Y   D   E   N   K   N   G   L   L   V   L   D
3408 aat gat tcc tac gat gaa aat aaa aac ggc ttg ctt gtt ctc gat
!
!   E   C   G   T   W   F   N   T   R   S   W   N   D   K   E
45
3453 gag tgc ggt act tgg ttt aat acc cgt tct tgg aat gat aag gaa
!
!   R   Q   P   I   I   D   W   F   L   H   A   R   K   L   G
3498 aga cag ccg att att gat tgg ttt cta cat gct cgt aaa tta gga
!
50
!   W   D   I   I   F   L   V   Q   D   L   S   I   V   D   K
3543 tgg gat att att ttt ctt gtt cag gac tta tct att gtt gat aaa
!
!   Q   A   R   S   A   L   A   E   H   V   V   Y   C   R   R
3588 cag gcg cgt tct gca tta gct gaa cat gtt gtt tat tgt cgt cgt
55
!
!   L   D   R   I   T   L   P   F   V   G   T   L   Y   S   L
3633 ctg gac aga att act tta cct ttt gtc ggt act tta tat tct ctt
!
!   I   T   G   S   K   M   P   L   P   K   L   H   V   G   V
60
3678 att act ggc tcg aaa atg cct ctg cct aaa tta cat gtt ggc gtt
!
!   V   K   Y   G   D   S   Q   L   S   P   T   V   E   R   W

```

```

3723 gtt aaa tat ggc gat tct caa tta agc cct act gtt gag cgt tgg
!
!   L   Y   T   G   K   N   L   Y   N   A   Y   D   T   K   Q
3768 ctt tat act ggt aag aat ttg tat aac gca tat gat act aaa cag
5
!   A   F   S   S   N   Y   D   S   G   V   Y   S   Y   L   T
3813 gct ttt tct agt aat tat gat tcc ggt gtt tat tct tat tta acg
!
!   P   Y   L   S   H   G   R   Y   F   K   P   L   N   L   G
10 3858 cct tat tta tca cac ggt cgg tat ttc aaa cca tta aat tta ggt
!
!   Q   K   M   K   L   T   K   I   Y   L   K   K   F   S   R
3903 cag aag atg aaa tta act aaa ata tat ttg aaa aag ttt tct cgc
!
!   V   L   C   L   A   I   G   F   A   S   A   F   T   Y   S
15 3948 gtt ctt tgt ctt gcg att gga ttt gca tca gca ttt aca tat agt
!
!   Y   I   T   Q   P   K   P   E   V   K   K   V   V   S   Q
20 3993 tat ata acc caa cct aag ccg gag gtt aaa aag gta gtc tct cag
!
!   T   Y   D   F   D   K   F   T   I   D   S   S   Q   R   L
4038 acc tat gat ttt gat aaa ttc act att gac tct tct cag cgt ctt
!
!   N   L   S   Y   R   Y   V   F   K   D   S   K   G   K   L
25 4083 aat cta agc tat cgc tat gtt ttc aag gat tct aag gga aaa TTA
!                                     PacI
!
!   I   N   S   D   D   L   Q   K   Q   G   Y   S   L   T   Y
30 4128 ATT AAt agc gac gat tta cag aag caa ggt tat tca ctc aca tat
!   PacI
!
!   i   I   D   L   C   T   V   S   I   K   K   G   N   S   N   E
!   iv                                     M1 K
35 4173 att gat tta tgt act gtt tcc att aaa aaa ggt aat tca aAT Gaa
!                                     Start
!
!   i   I   V   K   C   N   .End of I
!   iv   L3 L   N5 V   I7 N   F V10
40 4218 att gtt aaa tgt aat TAA T TTT GTT
!   IV continued.....
4243 ttc ttg atg ttt gtt tca tca tct tct ttt gct cag gta att gaa
atg
4291 aat aat tcg cct ctg cgc gat ttt gta act tgg tat tca aag caa
45 tca
4339 ggc gaa tcc gtt att gtt tct ccc gat gta aaa ggt act gtt act
gta
4387 tat tca tct gac gtt aaa cct gaa aat cta cgc aat ttc ttt att
tct
50 4435 gtt tta cgt gcA aat aat ttt gat atg gtA ggt tcT aAC cct tcc
atT
4483 att cag aag tat aat cca aac aat cag gat tat att gat gaa ttg
cca
4531 tca tct gat aat cag gaa tat gat gat aat tcc gct cct tct ggt
55 ggt
4579 ttc ttt gtt ccg caa aat gat aat gtt act caa act ttt aaa att
aat
4627 aac gtt ccg gca aag gat tta ata cga gtt gtc gaa ttg ttt gta
aag
60 4675 tct aat act tct aaa tcc tca aat gta tta tct att gac ggc tct
aat
4723 cta tta gtt gtt agt gcT cct aaa gat att tta gat aac ctt cct

```



```

!                                     RBS.?.?...
!       Start bla gene
!       6138 ATG agt att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg
gca ttt
5       6189 tgc ctt cct gtt ttt gct cac cca gaa acg ctg gtg aaa gta aaa
gat gct
!       6240 gaa gat cag ttg ggC gcA CTA GTg ggt tac atc gaa ctg gat ctc
aac agc
!                                     SpeI....
10      !                                     ApaLI & BssSI Removed
!       6291 ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg
atg agc
!       6342 act ttt aaa gtt ctg cta tgt GGC GcG Gta tta tcc cgt att gac
gcc ggg
15      6393 caa gaG CAA CTC GGT CGc cgC ATA cAC tat tct cag aat gac ttg
gtt gAG
!       BcgI.....
!       ScaI
!       6444 TAC Tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta
aga gaa
20      !       ScaI.
!       6495 tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac
tta ctt
!       6546 ctg aca aCG ATC Gga gga ccg aag gag cta acc gct ttt ttg cac
aac atg
25      !       PvuI....
!       6597 ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat
gaa gcc
!       6648 ata cca aac gac gag cgt gac acc acg atg cct gta gca atg Gca
aca acg
30      6699 tTG CGC Aaa cta tta act ggc gaa cta ctt act cta gct tcc cgg
caa caa
!       FspI....
!
35      6750 tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca ctt ctg
cgc tcg
!       6801 GCC ctt ccG GCT ggc tgg ttt att gct gat aaa tct gga gcc ggt
gag cgt
!       BglI.....
40      6852 gGG TCT Cgc ggt atc att gca gca ctg ggg cca gat ggt aag ccc
tcc cgt
!       BsaI....
!       6903 atc gta gtt atc tac acG ACg ggg aGT Cag gca act atg gat gaa
cga aat
45      !       AhdI.....
!       6954 aga cag atc gct gag ata ggt gcc tca ctg att aag cat tgg TAA
ctgt
!
!                                     stop
50      7003 cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt
taatttaaaa
!       7063 ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatccottaa
cgtgagtttt
!       7123 cgttccactg tacgtaagac cccc
!       7147 AAGCTT GTCGAC tgaa tggcgaatgg cgctttgcct
55      !       HindIII SalI..
!       (2/2) HincII
!       7183 ggtttccggc accagaagcg gtgcgggaaa gctggctgga gtgcgatctt
!
! Start of Fab-display cassette, the Fab DSR-A05, selected for
60      ! binding to a protein antigen.
!
!       7233 CCTGAcG CTCGAG

```

```

! xBsu36I XhoI..
!
! PlacZ promoter is in the following block
!
5 7246 cgcaacgc aattaatgtg agttagctca
7274 ctcattaggc accccaggct ttacacttta tgcttccggc tcgtatgttg
7324 tgtggaattg tgagcggata acaatttcac acaggaaaca gctatgacca
7374 tgattacgCC AagcttTGGa gccttttttt tggagatttt caac
!
! PflMI.....
10 ! Hind3. (there are 3)
! Gene iii signal sequence:
! 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
! M K K L L F A I P L V V P F Y
7418 gtg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat
15 !
! 16 17 18 Start light chain (L20:JK1)
! S H S A Q D I Q M T Q S P A
7463 tct cac aGT GCA Caa gac atc caq atg acc caq tct cca gcc
!
! ApaLI...
20 ! Sequence supplied by extender.....
!
! T L S L
7505 acc ctg tct ttg
!
! S P G E R A T L S C R A S Q G
25 7517 tct cca ggg gaa aga gcc acc ctc tcc tgc agg gcc agt cag Ggt
!
! V S S Y L A W Y Q Q K P G Q A
7562 gtt agc agc tac tta gcc tgg tac cag cag aaa cct ggc cag gct
30 !
! P R L L I Y D A S S R A T G I
7607 ccc agg ctc ctc atc tat gAt gca tcc aAc agg gcc act ggc atc
!
! P A R F S G S G P G T D F T L
35 7652 cca gCc agg ttc agt ggc agt ggg Cct ggg aca gac ttc act ctc
!
! T I S S L E P E D F A V Y Y C
7697 acc atc agc agC ctA gag cct gaa gat ttt gca gtT tat tac tgt
!
! Q Q R S W H P W T F G Q G T R
40 7742 cag cag CGt aAc tgg cat ccg tgg ACG TTC GGC CAA GGG ACC AAG
!
! V E I K R T V A A P S V F I F
7787 gtg gaa atc aaa cga act gtg gCT GCA Cca tct gtc ttc atc ttc
45 !
! BsgI....
!
! P P S D E Q L K S G T A S V V
7832 ccg cca tct gat gag cag ttg aaa tct gga act gcc tct gtt gtg
!
! C L L N N F Y P R E A K V Q W
50 7877 tgc ctg ctg aat aac ttc tat ccc aga gag gcc aaa gta cag tgg
!
! K V D N A L Q S G N S Q E S V
7922 aag gtg gat aac gcc ctc caa tcg ggt aac tcc cag gag agt gtc
55 !
! T E R D S K D S T Y S L S S T
7967 aca gag cgg gac agc aag gac agc acc tac agc ctc agc agc acc
!
! L T L S K A D Y E K H K V Y A
60 8012 ctg acG CTG AGC aaa gca gac tac gag aaa cac aaa gtc tac gcc
!
! EspI.....

```

```

!      C   E   V   T   H   Q   G   L   S   S   P   V   T   K   S
8057   tgc gaa gtc acc cat cag ggc ctG AGC TCg ccc gtc aca aag agc
!                                     SacI....
!
5      F   N   R   G   E   C   .   .
8102   ttc aac agg gga gag tgt taa taa
!
8126   GGCGCG CCaattctat ttcaaGGAGA cagtcata
!       AscI..... RBS2.
10
!       PelB signal sequence----- (22 codons)----->
!       1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
!       M  K  Y  L  L  P  T  A  A  A  G  L  L  L  L
8160   atg aaa tac cta ttg cct acg gca gcc gct gga ttg tta tta ctc
15
!       ...PelB signal-----> Start VH, FR1----->
!       16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
!       A  A  Q  P  A  M  A  E  V  Q  L  L  E  S  G
8205   gcG GCC cag ccG GCC atg gcc gaa gtt CAA TTG tta gag tct ggt
20      SfiI..... MfeI...
!       NcoI....
!
!       31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
!       G  G  L  V  Q  P  G  G  S  L  R  L  S  C  A
25      8250   ggc ggt ctt gtt cag cct ggt ggt tct tta cgt ctt tct tgc gct
!
!       ...FR1-----> CDR1-----> FR2-----
!
!       46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
!       A  S  G  F  T  F  S  T  Y  E  M  R  W  V  R
30      8295   gct TCC GGA ttc act ttc tct act tac gag atg cgt tgg gtt cgC
!       BspEI..
!       BstXI...
35
!       FR2-----> CDR2 -----
!
!       61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
!       Q  A  P  G  K  G  L  E  W  V  S  Y  I  A  P
8340   CAa gct ccT GGt aaa ggt ttg gag tgg gtt tct tat atc gct cct
40      BstXI.....
!
!       ...CDR2-----> FR3-----
!
!       76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
!       S  G  G  D  T  A  Y  A  D  S  V  K  G  R  F
45      8385   tct ggt ggc gat act gct tat gct gac tcc gtt aaa ggt cgc ttc
!
!       91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
!       T  I  S  R  D  N  S  K  N  T  L  Y  L  Q  M
50      8430   act atc TCT AGA gac aac tct aag aat act ctc tac ttg cag atg
!       XbaI...
!       Supplied by extender-----
!
!       -----FR3----->
55      106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
!       N  S  L  R  A  E  D  T  A  V  Y  Y  C  A  R
8475   aac acC TTA AGg gct gag gac act gca gtc tac tat tgt gcg agg
!       AflII...
!       from extender----->
60

```

```

! CDR3----->
FR4-->
! 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
! R L D G Y I S Y Y Y G M D V W
5 8520 agg ctc gat ggc tat att tcc tac tac tac ggt atg GAC GTC tgg
! AatII..
!
! 136 137 138 139 140 141 142 143 144 145
! G Q G T T V T V S S
10 8565 ggc caa ggg acc acG GTC ACC gtc tca agc
! BstEII...
!
! CH1 of IgG1----->
! A S T K G P S V F P L A P S S
15 8595 gcc tcc acc aag ggc cca tcg gtc ttc ccc ctg gca ccc tcc
! tcc
!
! K S T S G G T A A L G C L V K
20 8640 aag agc acc tct ggg ggc aca gcg gcc ctg ggc tgc ctg gtc
!
! D Y F P E P V T V S W N S G A
25 8685 gac tac ttc ccc gaa ccg gtg acg gtg tcg tgg aac tca ggc
! gcc
!
! L T S G V H T F P A V L Q S S
30 8730 ctg acc agc ggc gtc cac acc ttc ccg gct gtc cta cag tcc
! TCA
!
! Bsu36I....
!
! G L Y S L S S V V T V P S S S
35 8775 GGa ctc tac tcc ctc agc agc gta gtg acc gtg ccc tcc agc
! agc
! Bsu36I....
!
! L G T Q T Y I C N V N H K P S
40 8820 ttg ggc acc cag acc tac atc tgc aac gtg aat cac aag ccc
! agc
!
! N T K V D K K V E P K S C A A
45 8865 aac acc aag gtg gac aag aaa gtt gag ccc aaa tct tgt GCG
! GCC
!
! NotI.....
!
! A H H H H H H G A A E Q K L I
50 8910 GCa cat cat cat cac cat cac ggg gcc gca gaa caa aaa ctc
! atc
! ..NotI.... H6 tag..... Myc-
! Tag.....
!
! S E E D L N G A A q A S S A
55 8955 tca gaa gag gat ctg aat ggg gcc gca tag GCT AGC tct gct
! Myc-Tag.....
! ... NheI...
! Amber
!
! III'stump
!
60 ! Domain 3 of III -----
!
!

```



```

!       S   G   D   F   D   Y   E   K   M   A   N   A   N   K   G   A
8997 agt ggc gac ttc gac tac gag aaa atg gct aat gcc aac aaa GGC
GCC
!       tcc  t   t   t   t   t   a   g           a   c   t   t   g   g
5  t !W.T.
!
KasI...(2/4)
!
!       M   T   E   N   A   D   E   N   A   L   Q   S   D   A   K   G
10  9045 atG ACT GAG AAC GCT GAC GAG aat gct ttg caa agc gat gcc aag
ggt
!       c   a   t   c   t   a   c   g   c   a   g   t   c   t   c   t   a
c !W.T.
!
15  !       K   L   D   S   V   A   T   D   Y   G   A   A   I   D   G   F
9093 aag tta gac agc gTC GCG Acc gac tat GGC GCC gcc ATC GAc ggc
ttt
!       a   c   t   t   t   c   t   t   t   t   t   t   t   t   t   t
c !W.T.
!
20  !                               NruI....           KasI...(3/4)
!
!       I   G   D   V   S   G   L   A   N   G   N   G   A   T   G   D
9141 atc ggc gat gtc agt ggt ttG GCC Aac ggc aac gga gcc acc gga
gac
25  !       t   t   c   t   tcc  c   c   t   t   t   t   t   t   t   t   t
t !W.T.
!
!                               MscI....(3/3)
!
!       F   A   G   S   N   S   Q   M   A   Q   V   G   D   G   D   N
30  9189 ttc GCA GGT tcG AAT TCt cag atg gcC CAG GTT GGA GAT GGg gac
aac
!       t   t   c   t           c   a           t   a   c   t   c   t   t
t !W.T.
!
35  !       BspMI.. (2/2)           XcmI.....
!       EcoRI...
!
!       S   P   L   M   N   N   F   R   Q   Y   L   P   S   L   P   Q
9237 agt ccg ctt atg aac aac ttt aga cag tac ctt ccg tct ctt ccg
cag
40  !       tca  t   t   a           t   t   c   c   t   a   t   t   a   t   c   c   t
a !W.T.
!
!       S   V   E   C   R   P   F   V   F   S   A   G   K   P   Y   E
9285 agt gtc gag tgc cgt cca ttc gtt ttc tct gcc ggc aag cct tac
gag
45  !       tcg  t   a   t   c   t   t   c   t   agc  t   t   a   a   t
a !W.T.
!
!       F   S   I   D   C   D   K   I   N   L   F   R
50  9333 ttc aGC Atc gac TGC gat aag atc aat ctt ttC CGC
!       t   tct  t   t   t   c   a   a   c   t   a   c   t   !W.T.
!       BstAPI.....           SacII...
!                               End Domain 3
!
55  !       G   V   F   A   F   L   L   Y   V   A   T   F   M   Y   V   F
9369 GGc gtt ttc gct ttc ttg cta tac gtc gct act ttc atg tac gtt
ttc
!       t   c   t   g   t   c   t   t   a   t   t   c   c   t           t   a
t !W.T.
60  !       start transmembrane segment
!
!       S   T   F   A   N   I   L           R   N   K   E   S

```

—

! Table 37: DNA seq of w.t. M13 gene iii

```

!
!      1   2   3   4   5   6   7   8   9  10  11  12  13  14  15
!      fM   K   K   L   L   F   A   I   P   L   V   V   P   F   Y
5      1579  gtg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat
!      Signal sequence.....
!
!      16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
!      S   H   S   A   E   T   V   E   S   C   L   A   K   P   H
10     1624  tct cac tcc gct gaa act gtt gaa agt tgt tta gca aaa ccc cat
!      Signal sequence> Domain 1-----
!
!      31  32  33  34  35  36  37  38  39  40  41  42  43  44  45
!      T   E   N   S   F   T   N   V   W   K   D   D   K   T   L
15     1669  aca gaa aat tca ttt act aac gtc tgg aaa gac gac aaa act tta
!      Domain 1-----
!
!      46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
!      D   R   Y   A   N   Y   E   G   C   L   W   N   A   T   G
20     1714  gat cgt tac gct aac tat gag ggt tgt ctg tgG AAT GCt aca ggc
!                                     BsmI....
!      Domain 1-----
!
!      61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
!      V   V   V   C   T   G   D   E   T   Q   C   Y   G   T   W
25     1759  gtt gta gtt tgt act ggt gac gaa act cag tgt tac ggt aca tgg
!      Domain 1-----
!
!      76  77  78  79  80  81  82  83  84  85  86  87  88  89  90
!      V   P   I   G   L   A   I   P   E   N   E   G   G   G   S
30     1804  gtt cct att ggg ctt gct atc cct gaa aat gag ggt ggt ggc tct
!      Domain 1-----> Linker 1-----
!
!      91  92  93  94  95  96  97  98  99 100 101 102 103 104 105
!      E   G   G   G   S   E   G   G   G   S   E   G   G   G   T
35     1849  gag ggt ggc ggt tct gag ggt ggc ggt tct gag ggt ggc ggt act
!      Linker 1----->
!
!      106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
!      K   P   P   E   Y   G   D   T   P   I   P   G   Y   T   Y
40     1894  aaa cct cct gag tac ggt gat aca cct att ccg ggc tat act tat
!      Domain 2-----
!
!      121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
!      I   N   P   L   D   G   T   Y   P   P   G   T   E   Q   N
45     1939  atc aac cct ctc gac ggc act taT CCG CCt ggt act gag caa aac
!                                     EciI....
!      Domain 2-----
!
!      136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
!      P   A   N   P   N   P   S   L   E   E   S   Q   P   L   N
50     1984  ccc gct aat cct aat cct tct ctt GAG GAG tct cag cct ctt aat
!                                     BseRI..
!      Domain 2-----
!
!      151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
!      T   F   M   F   Q   N   N   R   F   R   N   R   Q   G   A
55     2029  act ttc atg ttt cag aat aat agg ttc cga aat agg cag ggc gca
!      Domain 2-----
!
!      166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
!      L   T   V   Y   T   G   T   V   T   Q   G   T   D   P   V
60

```

```

2074 tta act gtt tat acg ggc act gtt act caa ggc act gac ccc gtt
! Domain 2-----
!
!
5 ! 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
! K T Y Y Q Y T P V S S K A M Y
2119 aaa act tat tac cag tac act cct gta tca tca aaa gcc atg tat
! Domain 2-----
!
!
10 ! 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
! D A Y W N G K F R D C A F H S
2164 gac gct tac tgg aac ggt aaa ttC AGa gaC TGc gct ttc cat tct
! AlwNI.....
! Domain 2-----
!
!
15 ! 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225
! G F N E D P F V C E Y Q G Q S
2209 ggc ttt aat gaG GAT CCa ttc gtt tgt gaa tat caa ggc caa tcg
! BamHI...
! Domain 2-----
!
20 ! 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
! S D L P Q P P V N A G G G S G
2254 tct gac ctg cct caa cct cct gtc aat gct ggc ggc ggc tct ggt
! Domain 2-----> Linker 2-----
!
25 ! 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255
! G G S G G G S E G G G S E G G
2299 ggt ggt tct ggt ggc ggc tct gag ggt ggt ggc tct gag ggt ggc
! Linker 2-----
!
30 ! 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270
! G S E G G G S E G G G S G G G
2344 ggt tct gag ggt ggc ggc tct gag gga ggc ggt tcc ggt ggt ggc
! Linker 2-----
!
35 ! 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285
! S G S G D F D Y E K M A N A N
2389 tct ggt tcc ggt gat ttt gat tat gaa aag atg gca aac gct aat
! Linker 2> Domain 3-----
!
40 ! 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
! K G A M T E N A D E N A L Q S
2434 aag ggg gct atg acc gaa aat gcc gat gaa aac gcg cta cag tct
! Domain 3-----
!
45 ! 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315
! D A K G K L D S V A T D Y G A
2479 gac gct aaa ggc aaa ctt gat tct gtc gct act gat tac ggt gct
! Domain 3-----
!
50 ! 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330
! A I D G F I G D V S G L A N G
2524 gct atc gat ggt ttc att ggt gac gtt tcc ggc ctt gct aat ggt
! Domain 3-----
!
55 ! 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345
! N G A T G D F A G S N S Q M A
2569 aat ggt gct act ggt gat ttt gct ggc tct aat tcc caa atg gct
! Domain 3-----
!
60 !

```

```

!      346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
!      Q   V   G   D   G   D   N   S   P   L   M   N   N   F   R
2614   caa gtc ggt gac ggt gat aat tca cct tta atg aat aat ttc cgt
!      Domain 3-----
5
!      361 362 363 364 365 366 367 368 369 370 371 372 373 374 375
!      Q   Y   L   P   S   L   P   Q   S   V   E   C   R   P   F
2659   caa tat tta cct tcc ctc cct caa tcg gtt gaa tgt cgc cct ttt
!      Domain 3-----
10
!      376 377 378 379 380 381 382 383 384 385 386 387 388 389 390
!      V   F   S   A   G   K   P   Y   E   F   S   I   D   C   D
2704   gtc ttt agc gct ggt aaa cca tat gaa ttt tct att gat tgt gac
!      Domain 3-----
15
!      391 392 393 394 395 396 397 398 399 400 401 402 403 404 405
!      K   I   N   L   F   R   G   V   F   A   F   L   L   Y   V
2749   aaa ata aac tta ttc cgt ggt gtc ttt gcg ttt ctt tta tat gtt
!      Domain 3-----> Transmembrane segment-----
20
!      406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
!      A   T   F   M   Y   V   F   S   T   F   A   N   I   L   R
2794   gcc acc ttt atg tat gta ttt tct acg ttt gct aac ata ctg cgt
!      Transmembrane segment-----> ICA--
25
!      421 422 423 424 425
!      N   K   E   S   .
2839   aat aag gag tct taa ! 2853
!      ICA-----> ICA = intracellular anchor
30
!      ----- End of Table -----
-----

```

Table 38: Whole mature III anchor M13-III  
derived anchor with recoded DNA

```

!
!      1   2   3
5  !      A   A   A
!      1   GCG gcc gca
!      NotI.....
!
!      4   5   6   7   8   9  10  11  12  13  14  15  16  17
10 !      H   H   H   H   H   H   G   A   A   E   Q   K   L   I
!      10   cat cat cat cac cat cac ggg gcc gca gaa caa aaa ctc atc
!
!      18  19  20  21  22  23  24  25  26  27  28  29
15 !      S   E   E   D   L   N   G   A   A   .   A   S
!      52   tca gaa gag gat ctg aat ggg gcc gca Tag GCT AGC
!
!                                     NheI...
!
!      30  31  32  33  34  35  36      37  38  39
!      D   I   N   D   D   R   M      A   S   T
20 !      88   GAT ATC aac gat gat cgt atg   gct tct act
! (ON_G37bot) [RC] 5'-c aac gat gat cgt atg gcG CAt Gct gcc gag aca
g-3'
!      EcoRV..
!      Enterokinase cleavage site.
25 !
!      Start mature III (recoded) Domain 1 ---->
!      40  41  42  43
!      A   E   T   V
!      118   |gcC|gaG|acA|gtC|
30 !      t   a   t   t ! W.T.
!
!      44  45  46  47  48  49  50  51  52  53  54  55  56  57  58
!      E   S   C   L   A   K   P   H   T   E   N   S   F   T   N
!      130 |gaa|TCC|tgC|CTG|GCC|AaG|ccT|caC|acT|gaG|aat|AGT|ttC|aCA|Aat|
35 !      agt t t a a a c t a a tca t t c
! W.T.
!
!      MscI....
!
!      59  60  61  62  63  64  65  66  67  68  69  70  71  72  73
40 !      V   W   K   D   D   K   T   L   D   R   Y   A   N   Y   E
!      175 |gtg|TGG|aaG|gaT|gaT|aaG|acC|CtT|gAT|CGA|TaT|gcC|aaT|taC|gaA|
!      c a c c a t t a t c t c t g !
! W.T.
!
!      BspDI...
45 !
!      74  75  76  77  78  79  80  81  82  83  84  85  86  87  88
!      G   C   L   W   N   A   T   G   V   V   V   C   T   G   D
!      220 |ggC|tgC|TtA|tgg|aat|gcC|ACC|GGC|GtC|gtT|gtC|TGC|ACG|ggC|gaT|
!      t t c g t a t a t t t t c !
50 ! W.T.
!
!      SgrAI..... BsgI....
!
!      89  90  91  92  93  94  95  96  97  98  99  100 101 102 103
!      E   T   Q   C   Y   G   T   W   V   P   I   G   L   A   I
55 !      265 |gaG|acA|caA|tgC|taT|ggC|ACG|TGg|gtG|ccG|atA|gGC|TTA|GCC|atA|
!      a t g t c t a t t t g c t t c !
! W.T.
!
!      PmlI.... BlpI.....
!

```

```

!   Domain 1-----> Linker 1----->
!       104 105 106 107 108 109 110 111 112 113 114 115 116 117 118
!       P   E   N   E   G   G   G   S   E   G   G   G   S   E   G
!   310 |ccG|gaG|aaC|gaA|ggC|ggC|ggT|AGC|gaA|ggC|ggT|ggC|AGC|gaA|ggC|
5   !       t   a   t   g   t   t   c tct   g   t   c   t tct   g   t   !
W.T.
!
!       Linker 1-----> Domain 2----->
!       119 120 121 122 123 124 125 126 127 128 129 130 131 132 133
!       G   G   S   E   G   G   G   T   K   P   P   E   Y   G   D
10  !   355 |ggT|GGA|TCC|gaA|ggA|ggT|ggA|acC|aaG|ccG|ccG|gaA|taT|ggC|gaC|
!       c   t   t   g   t   c   t   t   a   t   t   g   c   t   t   !
W.T.
!
!       BamHI..(2/2)
15  !
!       134 135 136 137 138 139 140 141 142 143 144 145 146 147 148
!       T   P   I   P   G   Y   T   Y   I   N   P   L   D   G   T
!   400 |acT|ccG|atA|CCT|GGT|taC|acC|taC|atT|aaT|ccG|TtA|gaT|ggA|acC|
!       a   t   t   g   c   t   t   t   c   c   t c c   c   c   t   !
20  W.T.
!
!       SexAI....
!
!       149 150 151 152 153 154 155 156 157 158 159 160 161 162 163
!       Y   P   P   G   T   E   Q   N   P   A   N   P   N   P   S
25  !   445 |taC|ccT|ccG|ggC|acC|gaA|caG|aaT|ccT|gcC|aaC|ccG|aaC|ccA|AGC|
!       T   G   t   t   t   g   a   c   c   t   t   t   t   t tct   !
W.T.
!
!       HindIII...
30  !
!       164 165 166 167 168 169 170 171 172 173 174 175 176 177 178
!       L   E   E   S   Q   P   L   N   T   F   M   F   Q   N   N
!   490 |TTA|gaA|gaA|AGC|caA|ccG|TtA|aaC|acC|ttT|atg|ttC|caA|aaC|aaC|
!       c t   G   G tct   g   t c t   t   t   c           t   g   t   t   !
35  W.T.
!
!       HindIII.
!
!       179 180 181 182 183 184 185 186 187 188 189 190 191 192 193
!       R   F   R   N   R   Q   G   A   L   T   V   Y   T   G   T
40  !   535 |CgT|ttT|AgG|aaC|CgT|caA|gGT|GCT|CtT|acC|gTG|TAC|AcT|ggA|acC|
!       a g   c c a   t a g   g   g   a t a   t   t   t   g   c   t   !
W.T.
!
!       HgiAI...           BsrGI...
!
!       194 195 196 197 198 199 200 201 202 203 204 205 206 207 208
!       V   T   Q   G   T   D   P   V   K   T   Y   Y   Q   Y   T
45  !   580 |gtC|acC|caG|GGT|ACC|gaT|ccT|gtC|aaG|acC|taC|taT|caA|taT|acC|
!       t   t   a   c   t   c   c   t   a   t   t   c   g   c   t   !
W.T.
!
!       KpnI...
!
!       209 210 211 212 213 214 215 216 217 218 219 220 221 222 223
!       P   V   S   S   K   A   M   Y   D   A   Y   W   N   G   K
!   625 |ccG|gtC|TCG|AGT|aaG|gcT|atg|taC|gaT|gcC|taT|tgg|aaT|ggC|aaG|
55  !       t   a   a tca   a   c           t   c   t   c           c   t   a   !
W.T.
!
!       BsaI....
!       XhoI....
!
!       224 225 226 227 228 229 230 231 232 233 234 235 236 237 238
!       F   R   D   C   A   F   H   S   G   F   N   E   D   P   F
60  !   670 |ttT|CgT|gaT|tgT|gcC|ttT|caC|AGC|ggT|ttC|aaC|gaa|gac|CCT|ttT|

```

```

!           C A a   C   c   t   c   t t c t   c   t   t   G   T   a   c !
W.T.
!
!           239 240 241 242 243 244 245 246 247 248 249 250 251 252 253
5  !           V   C   E   Y   Q   G   Q   S   S   D   L   P   Q   P   P
      715 |gtC|tgC|gaG|taC|caG|ggT|caG|AGT|AGC|gaT|TtA|ccG|caG|ccA|CCG|
!           t   t   a   t   a   c   a t c g t c t   c c g   t   a   t   t !
W.T.
! DrdI.....
10 AgeI.....
!
!   Domain 2----->   Linker 2----->
!           254 255 256 257 258 259 260 261 262 263 264 265 266 267 268
!           V   N   A   G   G   G   S   G   G   G   S   G   G   G   S
15 760 |GTT|AAC|gcG|ggT|ggT|ggT|AGC|ggC|ggA|ggC|AGC|ggC|ggT|ggT|AGC|
!           c   t   t   c   c   c t c t   t   t   t t c t   t   c   c t c t
! W.T.
! AgeI.....
!           HpaI...
20 !           HincII.
!
!   Linker 2----->
Domain 3-->
!           269 270 271 272 273 274 275 276 277 278 279 280 281 282 283
25 !           E   G   G   G   S   E   G   G   G   S   G   G   G   S   G
      805 |gaA|ggC|ggA|ggT|AGC|gaA|ggA|ggT|ggC|AGC|ggA|ggC|ggT|AGC|ggC|
!           g   t   t   c t c t   g   t   c   t t c t   g   t   c t c t   t
! W.T.
!
!           -----Domain 3----->
!           284 285 286 287 288 289 290 291 292 293 294 295 296 297 298
!           S   G   D   F   D   Y   E   K   M   A   N   A   N   K   G
30 850 |AGT|ggC|gac|ttc|gac|tac|gag|aaa|atg|gct|aat|gcc|aac|aaa|GGC|
!           t c c   t   t   t   t   t   a   g           a   c   t   t   g   g !
! W.T.
35 KasI....
!
!           299 300 301 302 303 304 305 306 307 308 309 310 311 312 313
40 !           A   M   T   E   N   A   D   E   N   A   L   Q   S   D   A
      895 |GCC|atg|act|gag|aac|gct|gac|gaG|AAT|GCA|ctg|caa|agt|gat|gCC|
!           t           c   a   t   c   t   a   c   g   a   g t c t   c   t !
W.T.
! KasI....
45 StyI...
!           BsmI....
!
!           314 315 316 317 318 319 320 321 322 323 324 325 326 327 328
!           K   G   K   L   D   S   V   A   T   D   Y   G   A   A   I
50 940 |AAG|GGt|aag|tta|gac|agc|gTC|GCc|Aca|gac|tat|ggT|GCT|gcc|atc|
!           a   c   a c t   t t c t           t   t   t   c           t   !
W.T.
! StyI.....
!           PflFI.....
!
!           329 330 331 332 333 334 335 336 337 338 339 340 341 342 343
55 !           D   G   F   I   G   D   V   S   G   L   A   N   G   N   G
      985 |gac|ggc|ttt|atc|ggc|gat|gtc|agt|ggt|ctg|gct|aac|ggc|aac|gga|
!           t   t   c   t   t   c   t t c c   c c t           t   t   t   t !
W.T.
!

```



60



Table 40: Phage titers and enrichments of selections with  
a DY3F31-based human Fab library

	Input (total cfu)	Output (total cfu)	Output/input ratio
5 <b>R1-ox selected on phOx-BSA</b>	$4,5 \times 10^{12}$	$3,4 \times 10^5$	$7,5 \times 10^{-8}$
<b>R2-Strep selected on Strep-beads</b>	$9,2 \times 10^{12}$	$3 \times 10^8$	$3,3 \times 10^{-5}$

100574-103604

5 Table 41: Frequency of ELISA positives in  
DY3F31-based Fab libraries

	Anti-M13 HRP	9E10/RAM-HRP	Anti-CK/CL Gar-HRP
	R2-ox (with IPTG induction)	18/44	10/44
	R2-ox (without IPTG)	13/44	ND
10	R3-strep (with IPTG)	39/44	38/44
	R3-strep (without IPTG)	33/44	ND